AN EXAMINATION OF CONSTRUCT BIAS IN THE NARCISSISTIC PERSONALITY QUESTIONNAIRE FOR CHILDREN-REVISED (NPQC-R) ACROSS CULTURE, GENDER, AND AGE

By

Wei Yang Loke, Ed.S.

Submitted to the graduate degree program in the Department of Psychology and Research in Education and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Chairperson Patricia A. Lowe, Ph.D.

_____

Changming Duan, Ph.D.

_____

Amy N. Mendenhall, Ph.D.

_____

Tamara C. Mikinski, Ph.D.

_____

Matthew R. Reynolds, Ph.D.

Date Defended: April 22nd, 2015

The Dissertation Committee for Wei Yang Loke

certifies that this is the approved version of the following dissertation:

An Examination of Construct Bias in the Narcissistic Personality Questionnaire for Children-Revised (NPQC-R) across Culture, Gender, and Age

_____

Chairperson Patricia Lowe, Ph.D.

Date approved: May 3rd, 2015

Abstract

Construct bias in the Narcissistic Personality Questionnaire for Children – Revised (NPQC-R) was examined using the factor analytic and reliability methods among 701 Singapore and U.S. children and adolescents, ages 12 to 14 years. The NPQC-R is a 12-item measure that comprises of the Superiority and Exploitativeness subscales, and a Total scale. Results from the factor analytic method indicated that there was partial strong measurement invariance on the NPQC-R across culture (Singapore, U.S.) and gender (males, females), and strong measurement invariance on the NPQC-R across age (preadolescents, early adolescents). Also, a preponderance of evidence suggests that there is no construct bias across culture, gender, and age using the reliability methods based on an overlap in the 95% confidence interval of the internal consistency reliability estimates and tests of internal consistency reliability estimates using the Feldt technique between groups on the NPQC-R scale and subcale scores. Results from an examination of latent means indicated that U.S. students reported higher levels of Superiority compared to Singapore students, but no significant difference in the levels of Exploitativeness were found between groups. In addition, males and females had similar levels of Superiority, but males endorsed higher levels of Exploitativeness than females. Furthermore, preadolescents reported higher levels of Superiority compared to early adolescents, but no significant difference in the levels of Exploitativeness were found between the groups. Taken together, these results indicate no evidence of construct bias in the NPQC-R across culture, gender, and age. Limitations, future research directions, and practical implications of this study for practitioners (e.g., school psychologists) to become knowledgeable about the construct of narcissism, as measured by the NPQC-R is similar across culture, gender, and age, are discussed.

Acknowledgements

This journey would not have been complete without the support of many important people in my life. First, I would like to express my sincere gratitude to my advisor and dissertation chair, Trissy, for her encouragement throughout my six years of graduate school. Thank you for always providing me with sound advice and opportunities to grow personally and professionally. This dissertation would not have come to fruition if it were not for your gentle nudging and guidance throughout this process. I am also very grateful to Matt for his advice, and for always being very patient in explaining complex statistical methods to me. In addition, I greatly appreciate the constructive feedback from Drs. Changming Duan, Amy Mendenhall, and Tammy Mikinski in making this dissertation more polished.

I am very grateful to Rebecca for inspiring me to go into the field of school psychology. Your advice to apply to the University of Kansas for graduate school has afforded me the opportunity to learn from many great educators and researchers. In addition, I greatly appreciate your willingness to share the data that you had collected in Singapore.

I would like to express my heartfelt appreciation to my parents for their unwavering support throughout my life, especially during my graduate school endeavor. This exciting journey would not have been possible without your support. Thank you for believing in me and for instilling in me my thirst for knowledge. I would also like to thank my two sisters for their care and concern despite being many miles away.

A special thank you is reserved for my wonderful wife, Erika, for her unconditional love and support. Thank you for always cheering me on and for reminding me to celebrate the little successes in life. You have been very accommodating to my request to spend our weekends in

the library this past year. Words cannot express how grateful I am for all that you have done for me, and I look forward to spending more time with you during the weekends (not in the library).

Last, I would like to thank Him for giving me the strength to persevere through the ups and downs of graduate school, especially in the process of writing this dissertation.

List of Tables

List of Figures

TABLE OF CONTENTS

CHAPTER I

Introduction

The concept of narcissism has evolved over the past century. Its origins can be traced to the Greek mythology on Narcissus – a young hunter who fell in love with his own reflection in a mirrored pond, and subsequently died from unrequited love (Melville, 1986). This Greek mythology was popularized by Ovid in *Metamorphoses*. Later, the term narcissism was used by sexologists (e.g., Nacke) and psychoanalysts (e.g., Freud; Levy, Ellison, & Reynoso, 2011).

Contemporary conceptualizations of narcissism are dependent on whether narcissism is viewed as a clinical construct or a social-personality construct. Two prominent psychoanalysts, Kohut and Kernberg, brought the narcissism concept into the clinical realm. Indeed, Kernberg (1967) first described characteristics of patients with a "narcissistic personality structure", while Kohut (1968) postulated on the characteristics of patients with a "narcissistic personality disorder". The Diagnostic and Statistical Manual of Mental Disorders – Third Edition (DSM-III; American Psychiatric Association, 1980) subsequently included narcissistic personality disorder as one of the personality disorders. The Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (DSM-5; American Psychiatric Association, 2013) describes patients with narcissistic personality disorder as exhibiting characteristics relating to a grandiose sense of self, preoccupation with ideals, beliefs of being "special" or unique, needing overstated admiration by others, sense of entitlement, exploiting others, lack of empathy, feelings and beliefs of envy, and arrogant behaviors or attitudes. On the other hand, social and personality psychologists (e.g., Raskin & Hall, 1979), disagreed with clinicians that narcissism is a personality disorder that is diagnosed in clinical patients; instead, they view narcissism as a normal personality trait that is present in all individuals.

Owing to the study of narcissism by two different groups of researchers (i.e., clinicians and social and personality psychologists), narcissism has been conceptualized in a variety of ways. For example, the grandiose-vulnerable typology is prevalent among clinicians, whereas the overt-covert typology is used mostly by social and personality psychologists (Cain, Pincus, & Ansell, 2008; Wink, 1991). Another way that researchers have conceptualized narcissism is by organizing narcissism along a continuum from normal narcissism on one end of the spectrum to pathological narcissism on the other end of the spectrum. Normal narcissism reflects adaptive functioning, such as high self-esteem, while pathological narcissism reflects maladaptive functioning, such as excessive arrogance. However, some researchers (e.g., Ansell, 2005) view normal narcissism and pathological narcissism as separate dimensions. In addition, there is another group of researchers (e.g., Ang & Raine, 2009), who contend that narcissism is solely pathological due to its associations with maladaptive outcomes, such as aggression.

The extant literature on narcissism has mostly focused on adults; however, with the ushering in of the 21st century, a burgeoning interest in the study of narcissism in children and adolescents occurred. Narcissism in children and adolescents has been found to be associated with externalizing problems and mental health disorders. Studies have found positive relations between narcissism and aggression, especially proactive aggression (Ang & Raine, 2009; C. Barry, Grafeman, Adler, & Pickard, 2007; Fossati, Borroni, Eisenberg, & Maffei, 2010; Washburn, McMahon, King, Reinecke, & Silver, 2004). In addition, relations between narcissism and delinquency, and bullying, both traditional bullying and cyberbullying, have been found (Aalsma, Lapsley, & Flannery, 2006; Ang, Ong, Lim, & Lim, 2009; Ang, Tan, & Mansor, 2011; C. Barry, Grafeman et al., 2007; C. Barry, Pickard, & Ansel, 2009; C. Barry, & Wallace, 2010). Narcissism has also been found to be related to numerous mental health disorders, such as

conduct disorders (C. Barry, Frick, & Killian, 2003; T. Barry et al., 2007; C. Barry, & Wallace, 2010), oppositional defiant disorders (Frick, Bodin, & C. Barry, 2000), and attention-deficit/hyperactivity disorders (Frick et al., 2000). Taken together, relations between narcissism and negative functioning in children and adolescents warrant further examination into measures that examine narcissism in this population by mental health professionals with expertise in measurement, such as school psychologists.

Currently, the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988) is a widely used measure to assess adult subclinical narcissism, and it has served as the standard in the development of measures for children and adolescents. Numerous measures designed to assess childhood and adolescent narcissism were either downward extensions of the NPI, including the Narcissistic Personality Inventory for Children (NPIC; C. Barry et al., 2003) and the Narcissistic Personality Inventory – Juvenile Offender (NPI-JO; Calhoun, Glaser, Stefurak, & Bradshaw, 2000), or were guided theoretically by the NPI, including the Narcissistic Personality Questionnaire for Children (NPQC; Ang & Yusof, 2006) and the Narcissistic Personality Questionnaire for Children – Revised (NPQC-R; Ang & Raine, 2009). Of these measures, only the NPQC-R is a brief measure used to assess narcissistic features in children and adolescents. A couple of other brief measures of childhood and adolescent narcissism were also developed, including the Childhood Narcissism Scale (CNS; Thomaes, Stegge, Bushman, Olthof, & Denissen, 2008), and the narcissism scale on the Antisocial Process Screening Device (APSD; Frick et al., 2000). However, both the CNS and the APSD are unidimensional measures of narcissism. Most social and personality psychologists (e.g., Raskin & Terry, 1988), agree that narcissism is a multidimensional construct, with each dimension consisting of unique

characteristics that make up the narcissism construct. As such, the only brief, multidimensional measure of narcissism in children and adolescents currently is the NPQC-R.

The NPQC-R is a revision of an earlier measure, the NPQC. The original NPQC was developed with a sample of children and adolescents in Singapore, and consists of four factors: Superiority, Exploitativeness, Self-absorption, and Leadership. Ang and Raine (2009) undertook revisions to the NPQC due to low internal consistency reliability estimates for the scores on the Self-absorption and Leadership subscales. In addition, they asserted that although both the Self-absorption and Leadership subscales purport to measure adaptive narcissism, the inclusion of adaptive narcissism in a measure of narcissism is problematic as adaptive narcissistic characteristics are not core features of narcissism. Core features of narcissism that have been agreed on by researchers include grandiosity, entitlement, and exploitativeness (Dowson, 1992; Plakun, 1989). In addition, Ang and Raine (2009) contend that adaptive narcissism may not be necessarily adaptive. For example, in a longitudinal study conducted by C. Barry, Frick, Adler, and Grafeman (2007), adaptive narcissism was related to delinquency, after positive parenting was controlled. A revision to the NPQC, the NPQC-R, with a Singapore sample resulted in a 12-item measure that includes two dimensions of narcissism: Superiority and Exploitativeness. Items on the Superiority dimension tap an individual's inflated sense of self, while items on the Exploitativeness dimension tap characteristics of manipulativeness and taking advantage of others. Loke and Lowe (2014) subsequently validated the NPQC-R using a U.S. sample of children and adolescents in grades 4 through 8, and results supported a modified two-factor (Superiority and Exploitativeness) solution.

Test Bias. Test bias is one of the most important issues in the development of measures (Reynolds & Lowe, 2009). Bias is a statistical term used to determine whether differences on

measures assessing psychological traits or attributes of individuals are related to group membership (Reynolds & Lowe, 2009). Although bias is less frequently examined in psychological measures in comparison to cognitive tests, it is necessary to examine bias in these measures because a biased instrument may impact the interpretation of scores for individuals of different groups. In addition, it may influence important outcomes for different groups of individuals, such as misidentifying groups of individuals for intervention when they do not require it.

Although there are many aspects of test bias, construct bias has been viewed as a major concern when examining measures for bias (Reynolds, 1982). Construct bias relates to whether the measure assesses a psychological trait or attribute across groups with similar accuracy (Reynolds & Lowe, 2009). There is no one best way to test for construct bias; instead, it is recommended that multiple methods be used together to test for construct bias (Furr & Bacharach, 2013). Two methods to test for construct bias include reliability methods and factor analytic methods.

Pertaining to reliability methods, differences in internal consistency reliability estimates across groups of interest are examined to determine whether differences exist between these groups. An unbiased measure would generally display relatively similar internal consistency reliability estimates across groups. To test whether differences in internal consistency between groups exist, two techniques can be employed. First, the Feldt technique can be used (Feldt, 1969). This technique entails comparing a *F*-test statistic, which is obtained by dividing the larger error variance by the smaller error variance for groups, with the critical *F*-value from the *F*-distribution table. Should the *F*-test statistic not exceed the critical *F*-value, this would indicate that construct bias is not present across groups of interest (i.e., culture, gender, and age). In

addition, effect size estimates can be computed for each group difference. A negligible effect size (d = 0.2), according to Cohen (1988), indicates that differences between groups on the NPQC-R scale and subscale is minimal, and suggests that construct bias is not present between groups. Second, overlap in the 95% confidence intervals (CI) of internal consistency reliability estimates on the NPQC-R Total and subtests scores between groups can be examined. An overlap in the 95% CIs of the internal consistency reliability estimates between groups would indicate that construct bias is not present between the groups.

Factor analytic methods, such as exploratory factor analysis and confirmatory factor analysis, are used to determine whether the internal structure of a measure is similar or different across groups, respectively (Furr & Bacharach, 2013). Confirmatory factor analysis (CFA), and more specifically multi-group CFA, is used to determine whether there are differences in the latent factors of a measure across groups of interest (Reynolds & Lowe, 2009). Prior to performing multi-group CFAs, single-group CFAs are performed on each group to identify acceptable CFA models. Thereafter, multi-group CFAs are performed by testing for measurement invariance across groups of interest in a stepwise manner, by testing for configural invariance, weak invariance, and strong invariance in this order. This process begins by testing a least restricted model, with additional parameter constraints added to the preceding model at each step (Brown, 2015). Once strong factorial invariance has been established, comparisons between groups using latent means can occur.

Narcissism and Culture. The relationship between narcissism and culture has been equivocal (e.g., Fung, Gao, & Raine, 2010; Jonason, Li, & Czarna, 2013). Currently, only one study has examined differences in narcissism between the U.S. and Singapore. In Jonason et al.'s (2013) study with college students, students from the U.S. had higher levels of narcissism than

students from Singapore. Currently, no studies have examined differences in narcissism in children and adolescents in the U.S. and Singapore. However, a study of U.S. and Hong Kong students between the ages of 11 and 16 using a parent report found higher levels of narcissism in Hong Kong students than U.S. students (Fung et al., 2010). One reason has been proposed to explicate higher levels of narcissism in Asian cultures. Lasch (1978) suggested that capitalistic industrialization may contribute to an increase in narcissism. With higher gross national per capita in countries with predominantly Chinese populations (e.g., Hong Kong), individuals are more likely to prioritize personal goals over group goals, as they tend to pursue their own self-interests in order to accumulate wealth (Triandis, 1989). Taken together, although Singapore college students had lower levels of narcissism in comparison to U.S. college students, a study with children and adolescents in the U.S. and Hong Kong, an East Asian culture with a predominantly Chinese population, found higher levels of narcissism in Hong Kong students in comparison to students in the U.S. As such, further research needs to be conducted with children and adolescents in Singapore and the U.S. to better understand narcissism in these specific populations using a self-report measure.

Narcissism and Gender. In regard to the relation between narcissism and gender, studies have consistently found higher levels of narcissism in males in comparison to females (e.g., Twenge, 2006). This finding has been found across self- and parent-reports of narcissism, as well as across the U.S., and countries with predominantly Chinese populations (Thomaes et al., 2008; Zhou, Li, Zhang, & Zeng, 2012). Bond, Kwan, and Li (2000) suggested that males have been socialized to display gender roles, such as being more agentic and assertive. Twenge (1997) stated that these gender roles have been found to be related to narcissism. However, when gender differences were compared on the different dimensions of narcissism in children and adolescents,

results have been equivocal (Fossati et al., 2010; Washburn et al., 2004). For example, Fossati et al. (2010) found higher levels of exhibitionism in males than females, while Washburn et al. (2004) did not find any gender differences in exhibitionism. In addition, Washburn and colleagues (2004) found higher levels of adaptive narcissism in females in comparison to males, but Fossati et al. (2010) found higher levels of self-sufficiency and authority – two adaptive narcissistic characteristics, in males in comparison to females. Despite these findings, Fossati et al. (2010) and Washburn et al. (2004) agreed that males and females had similar levels of narcissistic exploitativeness. Taken together, despite narcissism appearing to be a male construct, gender differences on the various narcissistic dimensions have been inconsistent.

Narcissism and Age. Currently, no studies have examined construct bias in narcissism across ages, specifically between preadolescents and adolescents. However, some studies have examined the relationship between narcissism and age during childhood and adolescence (Carlson & Gjerde, 2009; Foster, Campbell, & Twenge, 2003). In Foster and colleagues' (2003) study of individuals between the ages of 8 and 83 years, young individuals between 8 and 14 years old were found to exhibit the highest level of narcissism among all other age groups, including adolescents between 15 and 19 years old. However, Carlson and Gjerde (2009) disputed Foster et al.'s (2003) findings; instead, they contend that the small sample size for the 8-14 year age group in Foster et al's (2003) study renders the need to interpret their findings with caution. Carlson and Gjerde (2009) found contrasting results in their longitudinal study, with adolescents having higher levels of narcissism at age 18 than when these individuals were 14 years of age. Research has found associations between narcissism and adolescent egocentrism (Aalsma et al., 2006; Alberts, Elkind, & Ginsberg, 2007). Adolescent egocentrism, which develops during adolescence, refers to distorted cognitive interpretations by the individual about

the self and others (Elkind, 1967). Specifically, studies have found consistent relationships between two aspects of adolescent egocentrism, i.e., personal uniqueness (being different from others), and omnipotence (having unlimited influence or power), and narcissism (Aalsma et al., 2006; Alberts et al., 2007). Although a few studies have looked into age differences in narcissism during childhood and adolescence, no study has explicitly examined differences in narcissism between preadolescents and early adolescents.

Statement of Purpose

The purpose of the present study is to examine construct bias in the NPQC-R, a measure of narcissism in children and adolescents, across culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents). Participants in this study were adolescents between the ages of 12 and 14 from Singapore and the U.S. To examine construct bias, both factor analysis and reliability methods were employed. Using the factor analysis method, tests of measurement invariance were performed on the data across culture, gender, and age, according to the stepwise procedure described by Brown (2015). This entails testing for configural invariance first, then weak factorial invariance and finally, strong factorial invariance. Construct bias was also examined by testing differences in internal consistency reliability estimates between groups of interest (i.e., culture, gender, and age) using Feldt's technique (Feldt, 1969; Reynolds & Lowe, 2009). In addition, the 95% confidence interval (CI) overlap in internal consistency reliability estimates between groups on the NPQC-R Total and its subscales was examined. If measurement invariance was found to be tenable across culture, gender, and/or age, latent means analyses would be performed to determine whether respective group differences exist on the NPQC-R.

Research Questions

1. Is there construct bias in the NPQC-R scores between groups using the factor analysis method?
    a. Does cultural bias exist on the NPQC-R between Singapore and U.S. children and adolescents?
    b. Does gender bias exist on the NPQC-R between male and female children and adolescents?
    c. Does age bias exist on the NPQC-R between preadolescents and early adolescents?
2. Is there construct bias in the NPQC-R scores between groups using the internal consistency reliability methods?
    a. Does cultural bias exist on the NPQC-R Singapore and U.S. children and adolescents?
    b. Does gender bias exist on the NPQC-R male and female children and adolescents?
    c. Does age bias exist on the NPQC-R between preadolescents and early adolescents?
3. Are there differences in narcissism between groups?
    a. Are there cultural differences in NPQC-R scores Singapore and U.S. children and adolescents?
    b. Are there gender differences in NPQC-R scores between male and female children and adolescents?

c. Are there age differences in NPQC-R scores between preadolescents and early adolescents?

Significance of the Study

This study has both theoretical and practical significance. This study contributes to the extant literature as this is the first study to examine construct bias in a narcissism measure designed for children and adolescents. Specifically, this is the first study to examine construct bias in the NPQC-R scores using multiple methods, such as factor analysis and reliability methods, across culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents). Construct bias is viewed as one of the most important aspects of test development (Reynolds & Lowe, 2009). If no evidence of construct bias was found with the NPQC-R scores across groups, then this would indicate that the NPQC-R assesses a similar underlying construct across groups, and it is plausible for future research to use the NPQC-R to develop theories across these respective groups (Reynolds, 1982).

The practical significance of this study pertains to practitioners in the U.S. being aware of the psychometric properties of the NPQC-R. By rejecting the notion that construct bias in the NPQC-R exists across culture, gender, and age, practitioners would be more aware that the construct of narcissism as measured by the NPQC-R is similar between Singapore and U.S. cultures, males and females, and preadolescents and early adolescents, and would feel more confident in using a measure that purports to measure the underlying narcissism construct across these groups. Also, practitioners would be more confident that an individual's score on the NPQC-R is not due to his or her group membership. As such, if an individual obtains an elevated score on the NPQC-R, practitioners may interpret this as the individual's "true score" and not due to his or her belonging to a certain group. In addition, these findings indicate that scores on

the NPQC-R for groups of interest are on the same metric, and suggests that comparisons can be made between groups using this measure. Taken together, should findings from this study indicate that cultural, gender, and age bias do not exist on the NPQC-R, then this would enable practitioners to interpret NPQC-R scores more accurately.

Summary

An increasing interest in childhood and adolescent narcissism has led to the development of measures to assess for narcissistic features in this population. The NPQC-R is a brief, multidimensional measure designed to assess narcissistic characteristics, specifically superiority and exploitativeness, in children and adolescents. This measure was first developed in Singapore, and has been validated with children and adolescents in the U.S. However, test (i.e., cultural, gender, and age) bias in the NPQC-R has yet to be examined. An important aspect of test bias is construct bias, and this refers to whether the measure assesses the same construct across groups. As such, this study seeks to examine construct bias in the NPQC-R across culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents). In addition, this study seeks to examine whether differences in narcissism across culture, gender, and age exists. Should findings indicate no construct bias in the NPQC-R across culture, gender, and age, findings from this study would demonstrate that the construct across cultures, genders, and ages are similar, and that it is appropriate for researchers to conduct further research and develop theories about narcissism across groups using the NPQC-R. In addition, findings may aid practitioners in comparing and interpreting scores between Singapore and U.S. cultures, males and females, and preadolescents and early adolescents, accurately, as they would be more confident that the scores on the NPQC-R are not due to individuals' group membership.

Chapter II

Review of the Literature

This chapter broadly reviews the extant literature on narcissism and test bias. The historical and contemporary accounts of narcissism are documented. The historical origins of narcissism are traced back to Greek mythology, while contemporary accounts of narcissism focuses on the psychoanalytic/clinical and the social/personality perspectives. As there is currently no "gold" standard in conceptualizing narcissism, numerous typologies of narcissism are described. In addition, measures to assess narcissism, especially measures of narcissism in children and adolescents, are noted. Specifically, the Narcissistic Personality Questionnaire for Children – Revised (NPQC-R; Ang & Raine, 2009), a multidimensional and short measure of maladaptive narcissism, is discussed in detail. Next, test bias is covered; in particular, construct bias and the methods to assess for construct bias are discussed. Thereafter, the relationships between narcissism and culture, gender, and age are presented.

Historical Myth of Narcissism

Narcissism has its origins in the Greek mythology on Narcissus. Although there are many variants of this Greek mythology, the version written by Ovid in his classic work titled *Metamorphoses* is considered the most popular (Melville, 1986). In Ovid's writings, Narcissus was a young hunter admired by many for his beauty. Although he had many suitors, both male and female, he rejected all of them, including a nymph called Echo. After being spurned by Narcissus, Echo spent her remaining days pining for Narcissus' love. Thereafter, one of Narcissus' other rejected suitors prayed to the Gods that Narcissus be punished. Nemesis, a Greek goddess of vengeance, answered his prayer by punishing Narcissus with unrequited love. One day, when Narcissus quenched his thirst in a mirrored pool, he fell in love with his own reflected image – an image that disintegrated whenever he touched the water. Narcissus could

not tear himself away from his image and pined for the love of his reflected image. After refusing food and rest, Narcissus subsequently died from unrequited love. While others were mourning his death, Narcissus' body disappeared, and a Narcissus flower lay in his place.

Clinical and Social Perspectives of Narcissism

The myth of narcissism was brought into the contemporary realm where it was theorized and studied by sexologists, e.g., Nacke, and psychoanalysts, e.g., Freud (Levy et al., 2011). Two prominent psychoanalysts, i.e., Kohut and Kernberg, subsequently influenced clinicians in their characterization of narcissism as a diagnosable pathological personality disorder that is present in clinical patients with mental disorders. Social and personality psychologists, however, disagreed with the clinical approach of narcissism; instead, they viewed narcissism as a normal personality trait that is present in all individuals. Owing to diverging viewpoints in the extant literature, narcissism is examined from two main perspectives: 1) psychoanalytic/clinical perspective, and 2) social/personality perspective (Cain et al., 2008).

Psychoanalytic/Clinical Perspective. The term *Narcissus-like* was first introduced by Ellis in 1898 to describe the "tendency for the sexual emotions to be lost and almost entirely absorbed in self-admiration" (Raskin & Terry, 1988, p.890). Nacke (1899) subsequently coined the term *Narcismus* to refer to autoeroticism as a perversion in which an individual treats one's body as a sexual object (Levy et al., 2011).

Freud adopted the term *Narcismus* and stated that autoeroticism was in fact a normal stage of early libidinal development, and that autoeroticism preceded the narcissistic phase of development, with narcissism being an intermediary phase between autoeroticism and true object love (Crockatt, 2006). In other words, as one matures, the individual seeks to develop true object love by directing his or her libidinal energy towards others; however, before a person can reach

this state, the individual goes through two other stages. In the autoeroticism phase, the individual's libidinal energy is released from his or her erogenous (or erotogenic) zones. As the individual gradually becomes more aware of the nurturance and love of one's mother, the person moves into the narcissistic stage and develops *primary narcissism*. Freud's (1914) seminal paper "On Narcissism" introduced two types of narcissism – primary and secondary. Primary (or normal) narcissism is present in all individuals, and occurs when the individual chooses the self as a libidinal object and views the self as omnipotent. This differs from secondary narcissism, which is associated with pathological conditions, and occurs when the libido withdraws from interest in objects in reality and develops an inflated form of omnipotence.

Freud's developmental approach was adopted by Kohut (1971) in his self-psychology theory of narcissism. Kohut (1971) conceptualized his idea of the 'self' as a system that organizes the individual's subjective experiences in relation to one's developmental needs (Banai, Mikulincer, & Shaver, 2005). As the self is the core of the individual's psychological development, healthy or pathological personality can be explained in reference to the self (Banai et al., 2005; Kohut, 1971/1977). Healthy expressions of narcissism, such as confidence and high self-esteem, arise from the development of a cohesive self-structure. The cohesive self-structure entails developing interpersonal relationships and connectedness to a group, developing and maintaining goal-setting ideals, as well as developing and maintaining a positive self-esteem Pathological narcissism results when difficulties in developing the cohesive self-structure arise. Kohut (1971/1977/1984) further outlined how pathological narcissism can result due to unmet developmental needs, i.e., needs for mirroring, idealization, and twinship (Kohut, 1971/1977/1984). A child will exhibit pathological narcissistic features when the caregiver fails to mirror responses of the child's grandiosity. In other words, if the caregiver is unable to admire

the child for his or her greatness and accomplishments, the child may develop pathological narcissistic characteristics. In addition, if a child's idealized image of his or her caregiver is destroyed, such that one is unable to admire and identify with his or her caregiver, the child is unable to exhibit healthy narcissistic characteristics. Taken together, cold and rejecting parents who are unavailable to mirror the grandiose self-image may result in an arrest in the child's development. In his subsequent paper, Kohut (1984) elaborated that a child who has an unmet need for twinship experiences, which is defined as not having an affiliation or similarity with or an interpersonal relationship with a group, such as one's family, may also be susceptible to pathological narcissism.

Similar to Kohut (1971), Kernberg (1975) expressed that pathological narcissism may occur due to expressions of narcissistic vulnerabilities that arise from early experiences. For example, when parents place excessive demands on their child and they demand that the child bring them glory without affirming their child's worth and value, the child may activate his/her unconscious ego defense mechanisms, so as to protect the grandiose self (Kernberg, 1975). One prominent defense mechanism is the devaluation of one's parents (Kernberg, 1976). When a child is rejected or not valued by one's parents, the individual would defensively withdraw from them, and inflate one's self-concept to minimize his/her perceived worthlessness (Russell, 1985). In addition, the child responds with aggression, hatred, and envy, to protect one's grandiose self. Apart from pathological narcissism, Kernberg (1975) posited that healthy narcissism could potentially result if the child develops positive relationships with one's parents.

Both Kernberg and Kohut were prominent in bringing narcissism into the clinical realm. Kernberg (1967/1970) was the first to describe the clinical characteristics of individuals with a "narcissistic personality structure", while Kohut (1968) took it one step further by formally

introducing the term "narcissistic personality disorder" and identified characteristics associated with this disorder (Levy et al., 2011). Some of these features include a belief of entitlement, exploitation of others, a lack of empathy, aggression, need for attention and admiration, a grandiose sense of self, and preoccupation with fantasies of success and power (Kernberg, 1975; Kohut, 1968). Kernberg's and Kohut's characteristics of narcissism greatly influenced the description of Narcissistic Personality Disorder in the Diagnostic and Statistical Manual of Mental Disorders – Third Edition (DSM-III; Ronningstam, 2005). Currently, in the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition (DSM-5; American Psychiatric Association, 2013), individuals who are diagnosed with a narcissistic personality disorder exhibit at least five of the following characteristics: a grandiose sense of self, preoccupation with infinite success, belief in one's uniqueness, sense of entitlement, need for admiration, lack of empathy, feelings of envy, conceited or arrogant attitudes and behaviors, and exploitation of others.

Social/Personality Perspective. Social and personality psychologists are interested in the study of narcissism as a personality trait that is present in the normal population. In other words, the focus of research by social and personality psychologists is on *subclinical* narcissism, rather than the study of narcissism as a personality disorder. The interest in narcissism emerged due to associations between narcissism and self-regulatory strategies to enhance one's self-esteem (Brown, Budzek, & Tamborski, 2009; Thomaes, Bushman, Orobio de Castro, & Stegge, 2009).

Self-regulatory models of narcissism. Several models have been developed to illustrate self-regulation among narcissistic individuals. The most influential model is Morf and Rhodewalt's (2001) *dynamic self-regulatory processing model of narcissism*. In this model, the narcissistic self is formed by the continuous interactions of the self with the social environment. Narcissistic individuals have a grandiose but fragile sense of self, and they seek to maintain and

defend their inflated sense of self through interpersonal strategies and intrapersonal processes. Interpersonal strategies entail manipulating others and self-enhancing oneself, so as to solicit positive responses and minimize negative feedback about the self from others. Intrapersonal processes that protect the grandiose sense of self include distorting or restructuring previous memories to reassure the self, or selectively attending to only positive feedback from others. In addition, the narcissistic individual's self-knowledge and social relationships can dynamically influence one's interpersonal strategies or intrapersonal processes. Indeed, the narcissistic individual's *self-knowledge* or view of the self and the social context may impact this self-regulatory process. A narcissistic individual's self-esteem may fluctuate as he/she tends to view one's traits differently in various social situations. As such, one strategy to enhance one's self would be to seek positive feedback from external sources, such as friends. Social relationships are highly important to narcissistic individuals. Specifically, narcissistic individuals gravitate towards individuals who can bolster their sense of self. However, when friends or acquaintances of the narcissistic individual view the person as imperfect or provide negative feedback that threatens the individual's fragile self, the narcissistic individual engages in self-aggrandizing interpersonal strategies or intrapersonal processes, such as insulting others to promote the self excessively, which in turn alienates one's friends (Buss & Chiodo, 1991). As such, narcissistic individuals are typically unable to maintain relationships with others because their goal in developing relationships is to construct a grandiose sense of self rather than social approval. Taken together, this results in an ongoing self-sustaining cycle to enhance and maintain one's grandiose sense of self, receive negative feedback or get rejected, and seek once again to enhance and maintain one's grandiose sense of self using self-regulatory mechanisms.

Other self-regulatory models have also been proposed to explain narcissism. Campbell, Brunell, and Finkel's (2006) *agency model of narcissism* extended the work of Morf and Rhodewalt (2001) and suggested that the five core traits of narcissistic individuals (i.e., a focus on agentic rather than communal concerns, approach-orientation, focus on obtaining self-esteem, entitlement, and an inflated view of self) mutually activate the use of interpersonal strategies and skills, such as self-promotion and charm, to generate narcissistic esteem (i.e., dominance-related self-esteem and excitement). In their amended *extended agency model of narcissism*, Campbell and Foster (2007) also mentioned that intrapsychic strategies, such as self-serving biases, are another set of regulatory strategies that narcissistic individuals use. Tracy and Robbins (2004) also proposed a model depicting how narcissistic individuals regulate their self-conscious emotions, such as shame and pride. Specifically, narcissistic individuals engage in evaluative processes to promote excessive pride and minimize shame. Narcissistic individuals may appraise positive events as being relevant to their identified goals, or attribute positive events to a stable aspect of the self. On the other hand, negative events are appraised as incongruent to their identified goals, and attributed to external causes or to unstable aspects of the self. For example, narcissistic individuals may be more prone to attribute high achievement to stable internal causes, while poor achievement is likely to be attributed to external causes. By doing so, the narcissistic individual maintains one's hubristic pride, while minimizing feelings of shame.

Typologies of Narcissism

Although narcissism has been studied across various fields (clinical psychology and social-personality psychology), there is currently no gold standard as to the definition of narcissism, which in turn results in different conceptualizations of the construct. Two prominent

typologies used to characterize narcissism are described below: 1) grandiose (overt) and vulnerable (covert) narcissism, and 2) normal and pathological narcissism.

Grandiose and vulnerable narcissism. The typology of grandiose and vulnerable narcissism has been widely discussed in the clinical literature, and has been elaborated upon by numerous researchers (e.g., Cain et al., 2008; Houlcroft, Bore, & Munro, 2012; Ronningstam, 2005). This grandiose-vulnerable distinction, which is used mostly by clinicians, parallels the overt-covert narcissism distinction, which is used frequently by social and personality psychologists (Wink, 1991). Despite this distinction, many researchers use these two typologies interchangeably.

Grandiose (overt) narcissism is characterized by outward expressions of achievement, sense of superiority or uniqueness, feelings of entitlement, attention-seeking behaviors and interpersonal exploitativeness; in comparison, vulnerable (covert) narcissism is characterized by feelings of anger, shame, and helplessness, with individuals being hypersensitive, anxious and insecure on the surface, but exhibiting grandiose behaviors once you get to know the individual better (Ronningstam, 2005). Support for the distinction between grandiose (overt) and vulnerable (covert) narcissism was found in two separate factor analytic studies (Rathvon & Holmstrom, 1996; Wink, 1991).

Support for this typology is also evidenced from the psychological outcomes of grandiose and vulnerable narcissism. Grandiose (overt) narcissism was found to be related to positive adjustment, such as self-esteem, happiness, and extraversion (Watson, Little, Sawrie, Biderman, 1992; Miller et al., 2011; Rose, 2002), whereas vulnerable (covert) narcissism was found to be associated with maladjustment outcomes, such as psychological distress, anxiety, and neuroticism (Houlcroft et al. 2012; Rathvon & Holmstrom, 1996; Miller et al., 2011). Rose

(2002) asserted that overt narcissistic individuals may obtain some positive benefits that their covert counterparts do not.

More recently, Cain et al. (2008) reviewed the literature on narcissism and identified two main themes of pathological narcissism, namely grandiosity-exhibitionism, and vulnerability-sensitivity-depletion. Of which, the authors contend that overt narcissism falls within the realm of grandiosity-exhibitionism, whereas covert narcissism falls within the vulnerability-sensitivity-depletion domain. Pincus et al. (2009) disagreed on the rigidity that an individual could only have either a grandiose or vulnerable narcissistic personality, and argued that both personalities may co-occur together, with the relative levels of grandiose and vulnerable self-states of narcissism determining which type of narcissistic personality the individual has at a given time. Extending this argument further, Pincus and Lukowitsky (2010) expressed that overt and covert narcissism should not be equated with grandiosity and vulnerability; instead, they should be viewed as different ways of expressing grandiosity and vulnerability. As such, overt and covert expressions of narcissism appear to be secondary classifications of grandiosity and vulnerability, and that it is plausible for individuals to exhibit overt and covert grandiosity, and overt and covert vulnerability.

Despite the grandiose-vulnerable narcissism distinction, there is still much emphasis on grandiose (overt) narcissism. Currently, most researchers (e.g., Ang & Raine, 2009; Rooningstam, 2005) agree that grandiose narcissistic characteristics make up the core of narcissism.

Normal and pathological narcissism. Narcissism has been conceptualized by some researchers (e.g., Cooper, 2005; Millon, Millon, Meagher, Grossman, & Ramnath, 2004; Ronningstam, 2005; Pincus & Lukowitsky, 2010) as being on a continuum from normal

narcissism to pathological narcissism. They contend that normal narcissism reflects *adaptive functioning* on one end of the spectrum, such as extraversion, self-confidence, and leadership qualities, while pathological narcissism reflects *maladaptive functioning* on the other end of the spectrum, such as excessive arrogance and feelings of superiority or uniqueness. Based on this conceptualization of narcissism, individuals with healthy narcissism would demonstrate adaptive characteristics, while individuals with unhealthy narcissism would exhibit maladaptive characteristics.

However, other researchers (e.g., Ansell, 2005; Pincus et al., 2009; Wink, 1991) assert that normal and pathological narcissism do not lie on the same continuum, but are in fact separate dimensions. Ansell (2005) performed a principal components analysis with measures of adult narcissism (e.g., NPI) and found support for adaptive and maladaptive narcissism as orthogonal dimensions of narcissism. The adaptive dimension consists of a continuum of characteristics ranging from high adaptive narcissism (e.g., high self-confidence, achievement motivation) to low adaptive narcissism (e.g., hypersensitivity, self-deprecation), while the maladaptive dimension consists of a continuum of characteristics ranging from high maladaptive narcissism (e.g., interpersonal exploitativeness, entitlement) to low maladaptive narcissism (e.g., empathetic understanding of interpersonal relationships, altruism). According to Ansell (2005), individuals who express healthy narcissism have high levels of adaptive narcissism and low levels of maladaptive narcissism, whereas individuals who express unhealthy narcissism have high levels of maladaptive narcissism and low levels of adaptive narcissism.

Yet, other researchers (e.g., Ang & Raine, 2009) do not view narcissism as adaptive; instead, they view narcissism as purely pathological or maladaptive due to frequent associations with maladaptive outcomes, such as conduct problems and delinquency. Indeed, in a longitudinal

study of 98 children and adolescents, adaptive narcissistic characteristics, such as self-sufficiency and authority, were initially not related to conduct problems, but after positive parenting was controlled, adaptive narcissism was related to conduct problems at two out of the three follow-ups (C. Barry, Grafeman et al., 2007). In other words, it is plausible that characteristics of self-sufficiency and authority may be predictive of delinquency. As such, adaptive narcissism does not appear to be necessarily adaptive, and that narcissism should be viewed from a purely pathological or maladaptive lens (Ang & Raine, 2009).

Assessment of Narcissism

Numerous measures have been developed to assess for narcissism, especially with adults. This section describes the different measures used to assess for adult narcissism, and childhood and adolescent narcissism. Specifically, the adult NPI is described in detail as most childhood and adolescent measures of narcissism, i.e., the NPIC, NPI-JO, NPQC-R, were either downward extensions or adapted from the NPI. With the NPQC-R being the focus of this dissertation, this section also elaborates on the process in which the NPQC-R was developed and its psychometric properties.

Measures of Narcissism in Adults. Measures of adult narcissism have focused on identifying adults with a narcissistic personality disorder (NPD). Clinical psychologists use a categorical approach to assess narcissism through criteria described in the DSM. In the DSM, NPD is characterized under the umbrella of personality disorders, with characteristics of NPD reflecting mostly grandiose features (Cain et al., 2008). Indeed, a confirmatory factor analysis on the criteria of NPD from the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition, Text Revision (DSM-IV-TR, American Psychiatric Association, 2000) supported a one-factor NPD model (Miller, Hoffman, Campbell, & Pilkonis, 2008).

Early self-report assessments of adults with NPD have frequently used the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1951) as a framework or template in the construction of measures. Two measures that adopted the Masculinity-Feminity (*Mf*) scale on the MMPI include Pepper and Strong's (1958) Ego-Sensitivity subscale and Serkownek's (1975) Narcissism-Hypersensitivity subscale. Subsequently, Ashby, Lee, and Duke (1979) developed the Narcissistic Personality Disorder Scale by identifying 19 items on the MMPI that were endorsed by individuals who were clinically diagnosed with NPD but were not endorsed by their non-clinical sample. Two other scales that were developed with items from the MMPI, but were adapted from the DSM-III NPD criteria were Morey, Waugh, and Blashfield's (1985) Narcissistic Personality Disorder Scale and Wink and Gough's (1990) Narcissism Scale. A main difference between these two scales was that Morey et al's (1985) scale was developed using a clinical sample, whereas Wink and Gough's (1990) scale was developed using a non-clinical sample. Other self-report measures of NPD have also been developed, e.g., the Narcissistic Personality subscale embedded within the Millon Clinical Multiaxial Inventory (MCMI; Millon, 1982), and the NPD scale within the Personality Diagnostic Questionnaire-4 (PDQ-4; Hyler, 1994).

Measures of subclinical narcissism have also been developed by social and personality psychologists. In contrast to the clinician's view of narcissism using a categorical approach, some social and personality psychologists (e.g., Thomaes et al., 2008) view narcissism along a continuum from less severe narcissistic behaviors (i.e., normal personality trait) to severe narcissistic behaviors (i.e., pathological narcissism; Foster & Campbell, 2007). Findings for this normal distribution of scores of narcissism were found in Foster and Campbell's (2007) non-clinical sample of 3,895 participants. However, most social and personality psychologists, e.g.,

Brown et al. (2009), disagree on the unidimensionality of narcissism, and assert that narcissism should be viewed from a multidimensional perspective. Because the construct of narcissism consists of different dimensions, and these dimensions have been found to have divergent relations with important variables, for example, narcissistic exploitativeness is the main facet of narcissism most strongly related to proactive aggression (Ang & Yusof, 2006), whereas narcissistic entitlement is most related to reactive aggression (Fossati et al., 2010), therefore, a dimensional approach is highly recommended as it would allow for a more detailed level of analysis of narcissism (Miller & Campbell, 2011).

Proponents of the distinction between grandiose (overt) and vulnerable (covert) narcissism have identified measures to assess both grandiose and vulnerable narcissism. Some measures used to assess vulnerable narcissism include the Hypersensitive Narcissism Scale (HNS; Hendin & Cheek, 1997), and the Narcissistic Vulnerability Scale (NVS; Bachar, Hadar, & Shalev, 2005), while one of the most frequently used measures used to assess grandiose narcissism is the Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988). With the core of narcissism consisting of grandiose narcissistic features, the NPI is therefore one of the most widely used measures of narcissism (Wright, Lukowitsky, Pincus, & Conroy, 2010).

***Narcissistic Personality Inventory*.** The NPI is the most prominent measure used to assess grandiose narcissistic features in non-clinical populations (Wright et al., 2010). Indeed, Cain and colleagues (2008) indicated that the NPI was the dominant instrument in social/personality research studies on narcissism since 1985; in fact, approximately 77% of the studies on narcissism utilized the NPI. The NPI was developed based on the Diagnostic and Statistical Manual of Mental Disorders-Third Edition (DSM-III; American Psychiatric Association, 1980) criteria of NPD (Raskin & Hall, 1979). It was first developed as a 54-item

measure using a sample of college students in the U.S. (Raskin & Hall, 1979), but Raskin and Terry (1988) shortened the original measure to a 40-item instrument, with each item consisting of dyadic statements about narcissistic features, and individuals were asked to select one statement that best describes them.

Limitations of the NPI have been documented in the extant literature. Two major limitations of the NPI include problems with the factor structure, and the item content within the NPI consisting of a confusing mix of adaptive and maladaptive characteristics. Regarding the former limitation, the factor structure of the NPI is equivocal (e.g., Emmons, 1984/1987; Raskin & Terry, 1988). Emmons (1984) first conducted a principal-components analysis on the NPI and extracted four factors: Exploitativeness/Entitlement, Leadership/Authority, Superiority/Arrogance, and Self-Absorption/Self-Admiration. The same four-factor solution of the NPI was confirmed in Emmons' (1987) exploratory factor analysis of the NPI. However, in Raskin and Terry's (1988) study, a principal component analysis performed on the 40-item NPI yielded seven factors: Authority, Self-Sufficiency, Superiority, Exhibitionism, Exploitativeness, Vanity, and Entitlement. Owing to discrepancies in the number of factors on the NPI from Emmons' (1984/1987) and Raskin and Terry's (1998) studies, Kubarych, Deary and Austin (2004) re-analyzed the NPI using principal components analysis and confirmatory factor analysis, and found that the NPI was better explained by a hierarchical model with three first-order factors and one higher-order general narcissism factor. The three first-order factors are Power, Exhibitionism, and Special Person. Corry, Merritt, Mrug, and Pamp (2008) extended the evaluation of the NPI by performing an exploratory and confirmatory factor analysis, and identified a two-factor solution, with the two factors being Leadership/Authority and Exhibitionism/Entitlement. Ackerman and colleagues (2011) claimed that Corry and colleagues'

(2008) study had underextracted factors due to their emphasis on scales with high internal consistency reliabilities, and they extracted three factors: Leadership/Authority, Grandiose/Exhibitionism, and Entitlement/Exploitativeness. Shortened versions of the NPI have also been developed to increase efficiency. For example, the NPI-13 (Gentile et al., 2013), which is a 13-item instrument that consists of three factors (Leadership/Authority, Grandiose/Exhibitionism, and Entitlement/Exploitativeness), and the NPI-16 (Ames, Rose, & Anderson, 2006), which is a 16-item unidimensional measure, were created. The second limitation of the NPI is it consists of a confusing mix of adaptive and maladaptive characteristics (Cain et al., 2008). Adaptive characteristics reflect features that include leadership and self-sufficiency, whereas maladaptive characteristics reflect features that include exploitativeness, entitlement and superiority (Ang & Raine, 2009). Due to poor internal consistency reliability estimates of its subscales, researchers have recommended the use of the total NPI score instead of its subscale scores (Tamborski & Brown, 2011). However, the total NPI score appears to measure healthy narcissism or adaptive aspects of narcissism (Pincus & Lukowitsky, 2010). For example, numerous studies (e.g., Brown et al., 2009; Sedikides, Rudich, Gregg, Kumashiro, & Rusbult, 2004) have found that the NPI scores were positively related to positive functioning, and were also negatively related to negative functioning, such as psychological distress, depression, sadness, anxiety, and neuroticism. As such, this alludes to the problematic nature of including adaptive and maladaptive features within the NPI, as the NPI, which was designed to measure grandiose (overt) narcissism, was found to be more related to adaptive than maladaptive functioning.

Measures of Narcissism in Children and Adolescents. Most instruments that assess narcissism in children have focused on self-reports (e.g., Narcissistic Personality Inventory for

Children). However, in situations when self-reports are used with young children, caution should be exercised as young children tend to hold an excessively positive self-view of themselves. As such, self-report assessments should only be used when children develop realistic perceptions of themselves (Marsh, Craven, & Debus, 1998). Having realistic perceptions of the self requires children to have the cognitive ability to evaluate their self-worth and to predict how others perceive their self-worth (Harter, 2006). Nicholls (1990) suggested that a child is capable of holding realistic views of himself or herself as early as age nine. Specific to narcissism, Thomaes et al. (2008) stated that it was even possible to obtain an accurate perception from children about their narcissistic features as early as age eight.

Despite the burgeoning interest in narcissism in children and adolescents, only a handful of measures to assess narcissism in children and adolescents have been developed. A couple of unidimensional measures used to assess childhood and adolescent narcissism include the Antisocial Process Screening Device (APSD; Frick et al., 2000) and the Childhood Narcissism Scale (CNS; Thomaes et al., 2008). Multidimensional measures of childhood and adolescent narcissism have also been developed. These measures were either downward extensions of the adult NPI, e.g., Narcissistic Personality Inventory for Children (NPIC; C. Barry et al., 2003) and the Narcissistic Personality Inventory – Juvenile Offender (NPI-JO; Calhoun et al., 2000), or were guided theoretically by the adult NPI, i.e., Narcissistic Personality Questionnaire for Children (NPQC; Ang & Yusof, 2006) and the NPQC-R (Ang & Raine, 2009). These measures are described in greater detail below.

Unidimensional Measures of Narcissism. The narcissism subscale on the APSD and the CNS are brief measures used to assess narcissism in children and adolescents (Frick et al., 2000; Thomaes et al., 2008). The APSD, which was first developed as the Psychopathy Screening

Device (PSD; Frick, O'Brien, Wootton, & McBurnett, 1994), consists of a self-report, and parent and teacher reports, and assesses various psychopathic behaviors in children and adolescents (Salekin, 2006). Although Frick et al.'s (1994) study identified two factors (callous/unemotional, and impulse control/conduct problems) with their clinic-referred sample , other researchers (e.g., Dadds, Fraser, Frost, & Hawes, 2005; Dong, Wu, & Waldman, 2014; Frick et al., 2000; Vitacco, Rogers, & Neumann, 2003) identified a three-factor (Callous-unemotional, Impulsivity, and Narcissism) solution in their samples of community, clinic-referred, and incarcerated children and adolescents. The narcissism subscale on the APSD consists of seven items that taps into grandiose (overt) narcissistic characteristics, and can be considered a unidimensional measure of narcissism (C. Barry & Ansel, 2011). Similarly, the CNS is a unidimensional measure designed to assess narcissistic features, such as a grandiose self-view and a negative interpersonal orientation, among children and adolescents. This 10-item measure was developed using a principal components analysis with 300 Dutch children and adolescents, and was subsequently validated using separate samples of Dutch and U.S. children and adolescents. Findings from Thomaes et al.'s (2008) study indicated that the one-factor CNS solution fit the data best. Despite the brevity of the narcissism subscale on the APSD and the CNS, unidimensional measures of narcissism may not be particularly useful, especially since different dimensions of narcissism relate differently to psychological functioning. For example, C. Barry et al. (2003) found that maladaptive narcissism, which consists of entitlement, exploitativeness, and exhibitionistic characteristics, were related to conduct problems, while adaptive narcissism, which consists of the authority and self-sufficiency characteristics, were not related to conduct problems. Further distinctions were found within the maladaptive narcissism realm, with one dimension of maladaptive narcissism, i.e., narcissistic exploitativeness, identified as the main

facet of narcissism that is most strongly related to aggression, while another dimension of maladaptive narcissism, i.e., narcissistic superiority, having negligible relations with aggression (Ang & Yusof, 2006),

Multidimensional Measures of Narcissism. The NPIC, NPI-JO, and NPQC-R are self-report measures that assess multiple dimensions of narcissism in children and adolescents. The NPIC and NPI-JO, which are downward extensions of the adult NPI, are first described below, and then the NPQC-R, which was developed using the adult NPI as a theoretical guide, is discussed.

Narcissistic Personality Inventory for Children (NPIC) and Narcissistic Personality Inventory – Juvenile Offender (NPI-JO). The NPIC was the first measure designed to specifically assess narcissism in a community sample of children and adolescents (C. Barry et al., 2003), whereas the NPI-JO was used to assess narcissism in juvenile offenders (Calhoun et al., 2000). Similar to the adult NPI, both measures required individuals to select a statement from two statements that best describes them. However, the response options for both measures differ. The NPI-JO retains the adult NPI's dichotomous (yes/no) response option, whereas the NPIC requires individuals to rate the statement that best describes the self on a 4-point Likert-type scale, so as to increase the variability of responses.

Both the NPIC and the NPI-JO consists of seven factors. On the NPIC, three items from the original NPI were removed due to low item-total correlations, resulting in a 37-item NPIC that retained the same seven dimensions of the NPI: Authority, Exhibitionism, Superiority, Entitlement, Exploitativeness, Self-Sufficiency, and Vanity (C. Barry et al., 2003). The factor structure of the NPIC was not subjected to analysis by the authors who developed the measure. However, Washburn et al. (2004) performed an exploratory factor analysis on the NPIC scores

among a sample of 233 students in the 5th through 8th grade, and found support for a three-(Adaptive Narcissism, Exhibitionism, and Exploitativeness) factor structure. In regard to the NPI-JO, results from a principal component analysis on 125 male juvenile offenders in the U.S. yielded fairly similar dimensions: Control/Vanity, Authority/Superiority, Exhibitionism, Leadership, Uniqueness, Need for Approval/Acknowledgement, and Exploitativeness. In addition, a total NPIC and a total NPI-JO score were obtained from a composite of their respective dimension scores (C. Barry et al., 2003; Calhoun et al., 2000).

Specific to the NPIC, C. Barry and colleagues (2003) further suggested that the measure consisted of adaptive and maladaptive narcissism dimensions. The authors contend that the adaptive narcissism dimension would consist of the Authority and Self-Sufficiency subscales due to their associations with positive psychological health, such as a need for achievement and assertiveness. In contrast, the maladaptive narcissism dimension would entail the Entitlement, Exhibitionism and Exploitativeness subscales due to their associations with poor psychological health, such as social maladjustment and poor impulse control.

Narcissistic Personality Questionnaire for Children (NPQC) and the Narcissistic Personality Questionnaire for Children – Revised (NPQC-R). The Narcissistic Personality Questionnaire for Children is the predecessor of the Narcissistic Personality Questionnaire for Children-Revised (NPQC-R). Based on the conceptual framework of the NPI and NPI-JO (Calhoun et al., 2000; Emmons, 1984/1987, Raskin & Terry, 1988), the NPQC was specifically designed to assess narcissistic characteristics among children and adolescents in the community. Initially, twenty-four items that covered the seven factors identified by Calhoun et al. (2000) were written; however, three items were subsequently removed after a review process by two independent reviewers. Unlike the NPI or NPI-JO, the individual was required to read only one

statement before rating oneself on a 5-point Likert-type scale, ranging from 1 (*not at all like me*) to 5 (*completely like me*). To explore the factor structure of the NPQC, data obtained from 370 children in grades 5 through 8 from Singapore were subjected to a principal-components analysis. Four factors were extracted from this analysis: Superiority, Exploitativeness, Self-absorption, and Leadership. Ang and Yusof (2006) then removed an additional three items due to their low item-total correlations. A second principal-component analysis followed with the 18-item NPQC. Results from this analysis yielded the same four factors identified from the first principal components analysis. The Superiority factor consists of items that describe the individual's inflated sense of self, while the Exploitativeness factor consists of items that assess characteristics of manipulativeness and taking advantage of others. The Self-absorption factor includes items that describe the individual's preoccupation with oneself and the need for acknowledgement from others, while the Leadership factor consists of items pertaining to being a leader and taking charge. These four factors are highly similar to the four factors described in Emmon's (1984/1987) NPI studies.

Ang and Raine (2009) emphasized a need for a revision of the NPQC due to several limitations arising from the instrument. First, both the Self-absorption and Leadership subscale scores on the NPQC had poor internal consistency reliabilities. Second, the authors contend that the inclusion of adaptive dimensions of narcissism in this instrument was problematic. They asserted that adaptive dimensions of narcissism, such as leadership and self-absorption, were not core features of narcissism. Instead, grandiosity is viewed as the most critical feature of narcissism (Ang & Raine, 2009; Plakun, 1989; Kernberg, 1975). In addition, some researchers (e.g., Dowson, 1992; Wink, 1996) have noted that reliable features of narcissism include characteristics of entitlement and exploitativeness. All three facets (i.e., grandiosity, entitlement,

and exploitativeness) are found on the maladaptive dimension of narcissism. Ang and Raine (2009) further questioned the items assessing adaptive characteristics on the adaptive dimension of narcissism. For example, they suggested that adaptive narcissism may not be really adaptive based on C. Barry, Frick et al.'s (2007) findings, as adaptive narcissism was related to delinquency after controlling for positive parenting.

Owing to the limitations of the adaptive dimensions on the NPQC, items on the Leadership and Self-absorption factors were dropped. As such, the NPQC-R consists of 12 items, with all items on the Superiority and Exploitativeness dimensions of the NPQC retained. Confirmatory factor analysis performed on the two-factor (Superiority and Exploitativeness) model and the four-factor (Superiority, Exploitativeness, Self-absorption, and Leadership) model using 479 7th and 9th grade students from the northern region of Singapore indicated that the two-factor model provided a better fit to the data. The authors further confirmed their findings that the two-factor model was a better fit to the data using a second sample of 470 7th and 9th grade students from the southern region of Singapore. As such, Ang and Raine (2009) asserted that the NPQC-R was a purer measure of maladaptive narcissism.

The NPQC-R was subsequently validated in the U.S. with a sample of students in grades 4 through 8 (Loke & Lowe, 2014). Results from the confirmatory factor analysis indicated that a modified two- (Superiority and Exploitativeness) factor model fit the data best. One item, Item 2, "If I ruled the world, it would be a better place" that loaded on the Exploitativeness factor in the Singapore sample, loaded on the Superiority factor in the U.S. sample. Ang and Yusof (2006) did note that this item had a high secondary item loading on the Superiority factor in their study and may have tapped onto characteristics associated with Superiority. Loke and Lowe (2014)

also contend that it was acceptable to respecify Ang and Yusof's (2006) model as the aforementioned item appeared to be more aligned with beliefs of an inflated sense of self.

Test Bias

Differences in test scores between groups may be due to an actual difference in the scores between groups, but it could also be due to the measure being biased towards one group (Furr & Bacharach, 2013). When bias in psychological tests exists, it presents difficulties in interpreting test scores for individuals of different groups accurately. Bias in psychological tests is concerned about whether there is systematic error in measuring a psychological trait or attribute across groups (Reynolds & Lowe, 2009). In other words, statistical analysis is used to identify whether differences in standardized tests of psychological traits or attributes in groups are due to membership in one group or another.

Examining bias in psychological measures is critical as tests can have important outcomes for individuals (Furr & Bacharach, 2013). Some important outcomes may include the use of test scores for diagnostic purposes or for determining if individuals need interventions. As such, it is necessary for the true score differences among individuals to not be due to group membership. For example, if a group of individuals were to score higher than another group on a test used to diagnose a disorder, it is necessary to ensure that this true score difference is not due to membership in one group vis-à-vis another group; if not, this would result in misdiagnosis and render inferences based on biased test scores as invalid (van de Vijver & Tanzer, 2004).

Test bias is one of the most important issues in test development, and it consists of three main sources, namely content bias, predictive bias, and construct bias (Reynolds & Lowe, 2009). Content bias refers to whether items on the test measure all aspects of the construct, whereas predictive bias refers to whether the test is able to predict future outcomes or behaviors.

Construct bias refers to whether a test measures similar psychological traits or attributes for different groups and with similar accuracy (Reynolds & Lowe, 2009).

Specifically, construct bias is examined in greater detail in this dissertation as it is deemed a major concern in test bias. Indeed, Reynolds (1982) indicated that if construct bias exists across groups, "much of the research of differential psychology of the present century [would need to] be discarded as confounded and major theories abandoned as primarily artifactual" (p. 200). In other words, if a test were to measure a psychological trait for one group but not for another group, and psychological research and theories were developed using this test for both groups, the research and theories would not be useful in explaining the psychological trait across all groups and should not be used.

Methods to Test for Construct Bias

There are a number of acceptable methods to assess for construct bias and it is recommended that these methods be used together (Furr & Bacharach, 2013). A couple of methods that have been suggested to examine construct bias include reliability and factor analytic methods.

Reliability Methods. A reliability method used to examine construct bias is the test of internal consistency reliability estimates between groups (Reynolds & Lowe, 2009). Internal consistency reliability estimates the degree to which all items on a subscale, or scale measure the same construct or dimension. An unbiased test should have very similar internal consistency reliability estimates across groups.

To statistically examine whether differences in internal consistency reliability estimates exist across different groups, the Feldt technique can be employed (Feldt, 1969). An *F*-test statistic is obtained by dividing the largest error variance by the smallest error variance for

groups on the same subscale or scale (see equation 1), and this *F*-test statistic is compared with the critical *F*-value from the *F*-distribution table. If the *F*-test statistic exceeds the critical *F*-value, this indicates that there is a significant difference between the two groups on the same scale or subscale, and suggests that there is construct bias across the groups. In addition, effect sizes for group differences can be computed. Effect size estimates can be interpreted as the percentage of non-overlap in scores between groups, with a smaller effect size indicating a smaller percentage of non-overlap between groups (Cohen, 1988). In other words, a negligible effect size, or an effect size estimate of less than 0.2, would indicate little to no difference in scores between groups, and would suggest that construct bias is not present between groups. The Feldt technique was performed in the present study with the NPQC-R Total and subscale scores across groups (i.e., gender, culture, age).

$$F = \frac{1 - alpha_1}{1 - alpha_2} \tag{1}$$

Another method that is used to test whether the internal consistency reliability estimates are similar between groups is to examine whether there is an overlap in the internal consistency reliability estimates' 95% confidence interval (CI) on the scales and subscales between groups. Should an overlap in the 95% CI between the groups occur, this would suggest that the internal consistency reliability estimates between the groups on the scales and subscales are relatively similar (Lowe, 2014). The 95% CI was computed for the NPQC-R scale and subscales for each group and comparisons were made across gender, culture, and age.

Studies using the NPQC-R using children and adolescents in Singapore have demonstrated adequate internal consistency reliability estimates for the scores of the NPQC-R

Total and subscales (e.g., Ang & Raine, 2009; Ang et al., 2011; Seah & Ang, 2008). Similar findings were reported in Loke and Lowe's (2014) U.S. sample of children and adolescents. However, no studies have reported internal consistency reliability estimates of the NPQC-R specific to boys and girls in either Singapore or the U.S, as well as across different age groups. Furthermore, no studies have used the internal consistency reliability method to test for construct bias in gender, culture, or age.

Factor Analytic Methods. Factor analytic methods are used to examine whether the internal structure of a test is similar or different between groups (Furr & Bacharach, 2013). The internal structure of a test pertains to the "pattern of correlations among items and/or the correlation between each item and the total test score" (Furr & Bacharach, 2013, p.6). Through factor analysis, latent constructs or factors are obtained from the correlations among observed variables (Reynolds & Keith, 2013). Exploratory factor analysis (EFA) and CFA are two types of factor analyses that are used to examine construct bias. EFA focuses on the degree of similarity of latent factors of the measure between groups, while CFA emphasizes the degree of differences on the latent factors of the measure between groups (Reynolds & Lowe, 2009).

Using multi-group CFA, measurement invariance between groups can be evaluated. Measurement invariance has been touted as one of the most important means of assessing test bias (Reynolds & Keith, 2013). Measurement invariance assumes that a test is unbiased, in that if individuals in different groups have similar latent traits or attributes, they should also have similar observed (item) scores (Meredith, 1993; Reynolds & Keith, 2013). As such, if no construct bias is present, measurement should be invariant across groups (Kline, 2011). In multi-group CFA, parameter constraints are added sequentially to multi-group models, with the goal

being to obtain a restricted and parsimonious model that is equivalent between groups (Reynolds & Keith, 2013).

To assess for measurement invariance, configural invariance, which is the least restricted model, needs to be established before other models with added constraints are tested. In a stepwise manner from the least restricted to the most restricted model are configural invariance, weak factorial invariance, and strong factorial invariance models (Meredith, 1993).

When configural invariance, or "equality of form", is established, this indicates that the measurement model is similar across groups, such that the number of factors, and the pattern of factor-indicator (latent factor-observed variable) correspondence across groups are similar (Brown, 2015; Kline, 2011). Once configural invariance has been established, weak factorial invariance can be tested.

Weak factorial invariance is also referred to as "metric invariance" or "equality of factor loadings". In this model, an additional constraint to equate unstandardized factor loadings across groups is placed (Kline, 2011). If weak factorial invariance is found, this suggests relatively equal factor loadings across groups, such that scores on the manifest scale have similar *meaning* in each group (Steinmetz, Schmidt, Tina-Booh, Wieczorek, & Schwartz, 2004). However, it has to be noted that weak factorial invariance only concludes that the factor loadings in one group are proportionally similar to the corresponding loadings in another group, and that conclusions that similar observed scores on an indicator at a specified level of the latent factor for each group cannot be drawn (Bontempo & Hofer, 2006; Brown, 2015). For the latter conclusion to be made, it is necessary that scaling bias be examined with a more restrictive model.

Once weak factorial invariance is established, a more restrictive model, i.e. strong factorial invariance, can be tested. Strong factorial invariance is also referred to as "scalar

invariance" or "equality of indicator (e.g., item) intercepts". In this model, an additional constraint on top of the weak factorial invariance model is imposed, with the indicator's intercepts (or thresholds) now constrained across groups, such that the latent factor of each group is now on the same scale (Bontempo & Hofer, 2006). This alludes to the latent factor being able to explain as much group differences in the observed indicators as possible (Meredith, 1993).

Taken together, the factor structure of a measure in each group is compared to identify if any differences exist in the factor structure. Should results find differences in the factor structure between groups, we would conclude that the internal structure for the groups are different, and that there is bias in test scores in the psychological measure being used.

Measurement Invariance Studies and the NPQC-R. Measurement invariance studies have been conducted with the NPQC-R. In Ang and Raine's (2009) sample of 479 Singapore adolescents, ages 12 - 17 years, the authors demonstrated that the NPQC-R scores were invariant between males and females, and that the NPQC-R scores were invariant across students in seventh grade and students in ninth grade. Findings from Loke and Lowe's (2014) study of 268 U.S. adolescents also indicated that the NPQC-R scores were invariant across males and females. In addition, the authors found that the NPQC-R scores were invariant across students in grades 4-6 and 7-8.

Testing for Group Differences. Relatedly, measurement invariance is an important condition that has to be fulfilled before groups can be compared. Indeed, researchers have asserted that it is only when strong factorial invariance is established that factor means can be compared (Brown, 2015; Reynolds & Keith, 2013). Without measurement invariance to indicate that the measure is unbiased across groups before testing for group differences on a

psychological attribute or trait, results may not accurately reflect the presence of, or lack of, difference among groups on the latent variable (Borsboom, 2006).

Most studies that have examined differences among groups have relied on examining differences in observed scores. Indeed, in narcissism research in children and adolescents, this method has been widely used to examine group differences (e.g., Thomaes et al., 2008). However, Sass (2011) stated that inferences drawn from comparing means in observed variables may differ from inferences drawn from comparing latent means. Comparing means of latent variables has an advantage over the mean differences of observed variables approach as latent variables exclude random measurement errors related to the latent variable; in contrast, observed scores consist of both systematic and random error components (Dimitrov, 2006). Thus, group differences that are found using the latent means comparison approach can be viewed with greater confidence as it rules out other possible reasons that may confound the results.

To compare latent means, Reynolds and Keith (2013) recommended that the output of the strong factorial invariance model be examined where the statistical significant tests for latent means are reported. If a statistically significant result is found and the latent mean is positive, then this would indicate that the non-reference group scored higher on the variable of interest than the reference group. On the other hand, if a statistically significant result is found and the latent mean is negative, then this would indicate that the reference group scored higher than the non-referenced group. Currently, there is no study within the field of childhood and adolescent narcissism that has examined cultural, gender, or age differences using the latent means comparison method.

Narcissism and Culture

Although behavioral genetics using twin studies have found that heredity accounts for approximately 60% of the variability in narcissism scores, environmental factors, such as the cultural environment, are important contributory factors of narcissism (Livesley, Jang, Jackson, & Vernon, 1996; Vernon, Villani, Vickers, & Harris, 2008). This section of the dissertation describes narcissism across cultures; specifically, narcissism levels were compared between U.S. and Asian countries with predominantly Chinese populations.

Empirical studies comparing narcissism between the U.S. and Asian countries with predominantly Chinese populations have been conducted. However, equivocal results have been reported (e.g., Cai & Gries, 2013; Fukunishi et al., 1996).

In an online study of 718 individuals between the ages of 11 and 66 from the U.S. and China, Cai and Gries (2013) found higher levels of narcissism in their sample of children and adults from the U.S. as compared to children and adolescents from China. In Jonason and colleagues' (2013) study of 626 college students, students from the U.S. and Poland had higher levels of narcissism than students from Singapore. It is plausible that higher levels of narcissism in the U.S. as compared to China or Singapore could be because of elevated levels of individualism in the U.S. as compared to these Asian countries (Hofstede, 1991). In fact, the late 1970s in the U.S. has been described as the "me" generation and the age of narcissism, with individuals forsaking common social goals for the continuous pursuit of the self (Kanfer, 1979). Indeed, Twenge, Konrath, Foster, Campbell, and Bushman (2008) found that narcissism levels, as measured with the Narcissistic Personality Inventory (NPI), among college students had increased by 30% in the U.S between 1979 and 2006.

In contrast, in Fukunishi and colleagues' (1996) study of 121 U.S. college students, 119 Japanese college students, and 70 Chinese college students, the authors found higher levels of narcissism on self-reports of narcissistic features using the NPI by college students from China than college students from the U.S. and Japan. Similarly, in Kwan, Kuang, and Hui's (2009) study of 116 college students in China and 131 college students in the U.S., their sample of Chinese college students had higher levels of narcissism, as assessed using the NPI, than college students from the U.S. This finding also holds true for children and adolescents. In a parent-report of narcissistic features in a large sample of children and adolescents, ages 11 to 16, using the narcissism scale on the APSD, higher levels of narcissism were found among students from Hong Kong than students from the U.S. (Fung et al., 2010). One reason to explicate higher levels of narcissism in Asian countries with predominantly Chinese populations has been postulated. As a culture becomes more complex, it becomes more individualistic (Triandis, 2001). One measure of cultural complexity is the gross national product per capita. Hofstede (1991) found that affluent countries are more likely to be individualistic, while poorer countries are more likely to be collectivistic. In wealthy countries, individuals prioritize personal goals over goals of the group, such that their pursuit of self-interest may result in accumulation of wealth (Triandis, 1989). As such, it appears that capitalism may shift towards an emphasis of personal goals over social and group goals (Kwan et al., 2009). Indeed, Lasch (1978) asserted that capitalistic industrialization contributes to a rise in narcissism.

Narcissism and Gender

Adult narcissism research has indicated that males have higher clinical and subclinical narcissism levels than females (e.g., Klonsky, Jane, Turkheimer, & Oltmanns, 2002; Twenge, 2006). In Bond et al.'s (2000) study of 235 college students from Hong Kong, they asserted that

males have elevated levels of narcissism as compared to females because males have been socialized to adhere to gender roles that include being more assertive and agentic, with the latter term referring to one's capacity to intentionally make choices to influence the self and one's environment, and reflects characteristics such as the need for power and achievement (Bandura, 2002). Both of these gender roles have been found to be related to narcissism (e.g., Twenge, 1997). Indeed, college men who behaved consistently with their gender role, i.e., men who were more masculine, were found to exhibit more narcissistic characteristics (Klonsky et al., 2002). Furthermore, Tschanz, Morf, and Turner (1998) found that narcissistic exploitativeness/entitlement features were more prevalent in males than females in their large study of 2,089 predominantly White college students in the U.S. Tsanchz et al. (1998) alluded that females tend to display less exploitativeness/entitlement characteristics due to the need to present desirable sex-role expectations, such as being gentle, warm, and sensitive. However, contradictory findings were found in Klonsky and colleagues' (2002) study of 665 predominantly White college students in the U.S., as self- and peer-reports of females who behaved consistently with their gender role, i.e., females who were more feminine, were rated as exhibiting more narcissistic characteristics.

Studies on narcissism among children and adolescents have found males to have higher narcissism levels than females (T. Barry et al., 2007; Fosatti et al., 2010; Frick et al., 2000; Kerig & Stellwagen, 2010; Ojanen, Findley, & Fuller, 2012; Zhou et al., 2012). For example, Frick and colleagues (2000) found that boys scored higher than girls on parent and teacher ratings of narcissism, whereas Thomaes et al. (2008) found that boys scored higher on self-report ratings of narcissism compared to girls. Similar findings were also reported in collectivistic cultures, such as China and Hong Kong. For example, Zhou and colleagues (2012) reported that junior high

school boys in China had higher narcissism scores than girls. Moreover, Fung and colleagues (2010) found similar results for parent report of narcissism in children from Hong Kong between the ages of 11 and 16 years. However, when narcissism was broken down into its various components, varying results were found. For example, in Fossati et al.'s (2010) study, no differences were found between Italian males and females on narcissistic characteristics of superiority, exploitativeness, and entitlement. In addition, although males scored higher on exhibitionism and the two adaptive narcissistic features, i.e., self-sufficiency and authority, when compared to females, females had higher levels of vanity compared to males. Similarly, Washburn et al's (2004) study of 233 U.S. adolescents found that females had significantly higher levels of adaptive narcissism than males; however, no differences in exhibitionism and exploitative characteristics were found between male and female adolescents. As such, although narcissism may appear to be a male-dominated construct, closer inspection of the findings suggest equivocal findings when narcissistic features that make up this construct were examined.

Narcissism and Age

This section describes age-related changes in narcissism during childhood and adolescence. Preadolescents are defined as individuals who are between 10 and 12 years, while individuals who are between the ages of 13 and 15 years are considered early adolescents (Emond, Ormel, Veenstra, & Oldehinkel, 2007; Kvaraceus, 1944).

Narcissism has been found to decline with age (Wilson & Sibley, 2011). In two separate samples of New Zealand participants, Wilson and Sibley's (2011) cross-sectional study documented a negative relationship between age and narcissism, with individuals between 20 and 30 years showing higher levels of narcissism as compared to other older age groups. It has to be noted that both samples were based on adult responses, with the first sample comprising of

6,507 participants with an average age of 47.80 years, and the second sample comprising of 2,525 participants with an average age of 29.57 years.

Currently, only a few studies (Foster et al., 2003; Carlson & Gjerde, 2009) have examined age-related changes in narcissism among adolescents, and results from these studies have been equivocal. In Foster and colleagues' (2003) cross-sectional study of participants between 8 and 83 years, they found that younger people endorsed more narcissistic characteristics on the NPI than older people. Closer inspection of their findings indicated that individuals between 8 and 14 years old had the highest level of narcissism among all age groups. Carlson and Gjerde (2009), however, stated that Foster et al's (2003) finding has to be interpreted with caution as the sample size of individuals between 8 and 14 years in their study was relatively small (*n*=45) as compared to other age groups, e.g., 15-19 (*n*=1219) or 20-24 (*n*=1029) age groups. When examining differences in narcissism during adolescence, Carlson and Gjerde's (2009) longitudinal study of 103 U.S. participants found results contrary to findings by Foster et al. (2003). In Carlson and Gjerde's (2009) study, students were assessed at one time point during early adolescence and middle adolescence, and results indicated that individuals reported higher levels of narcissism when they were at age 18 as compared to when they were 14 years old.

There are currently no studies that have specifically examined differences in narcissism between preadolescents and early adolescents. As children enter the adolescence phase, they develop adolescent egocentrism, which is defined as a cognitive distortion about how an individual interprets the attitudes of the self and others (Elkind, 1967). Adolescent egocentrism consists of two related but distinct concepts, i.e., imaginary audience and personal fables (Lapsley, FitzGerald, Rice, & Jackson, 1989). Imaginary audience refers to the presumption by

the adolescent that everyone is as preoccupied as him/her about his/her appearance and behavior, while personal fables consist of characteristics that are related to personal uniqueness (i.e., perception of being different from others and no one understanding the adolescent), omnipotence (i.e., the perception of having unlimited influence or power), and invulnerability (i.e., the perception of being incapable of being harmed or injured.

Alberts and colleagues (2007) suggested that adolescent egocentrism is intricately linked to the development of personality. Specifically, Lapsley et al. (1989) found that adolescent egocentrism characteristics of imaginary audience and personal fables were closely related to narcissism in their sample of 169 students in 6th, 8th, 10th, and 12th grade. In addition, another study conducted by Aalsma et al. (2006) found that two aspects of personal fables, i.e., personal uniqueness and omnipotence, were associated with narcissism, albeit invulnerability was not associated with narcissism.

Summary

The study of narcissism has evolved over the years from a clinical construct studied solely by psychoanalysts and clinicians to a construct that is viewed as a normal personality trait and is studied by social and personality psychologists. Although extensive research in narcissism has been conducted with adults, the study of narcissism in children and adolescents has been substantially less, but has increased with the turn of the 21st century.

With increasing interest in childhood and adolescent narcissism, measures to assess narcissistic features in this population have emerged. The NPQC-R is a brief, multidimensional measure designed to assess narcissistic features in children and adolescents (Ang & Raine, 2009). This measure is a revision of the NPQC, and was developed using a sample of Singapore

students. The NPQC-R consists of two main characteristics of narcissism, i.e., superiority and exploitativeness, and is viewed as a purer measure of maladaptive narcissism.

Test bias is an important concept in the development of psychological measures, such as the NPQC-R. A test that is biased towards one group when compared to another group would make accurate interpretation of test results difficult (van de Vijver & Tanzer, 2004). A major concern of test bias is construct bias, and construct bias can be assessed by using reliability methods, such as comparing internal consistency reliability estimates across groups, and using factor analytic methods, such as performing measurement invariance across groups (Reynolds & Lowe, 2009).

Group differences on variables, such as culture, gender, and age are better explained by comparing latent means across groups. Studies have typically compared sample means across groups to determine if differences in narcissism occur between groups. However, inferences obtained by comparing latent means between groups are more advantageous as they exclude random error. As such, this dissertation seeks to examine whether latent means are different across culture, gender, and age in childhood and adolescent narcissism.

Pertaining to cultural differences, specific comparisons on narcissism between the U.S. and cultures with predominantly Chinese populations have been conducted, with equivocal results found (e.g., Jonason et al., 2013; Fung et al., 2010). Recently, a cross-cultural study of narcissism between Singapore and U.S. college students found higher levels of narcissism among U.S. students than Singapore students (Jonason et al., 2013). Only one study has examined childhood and adolescent narcissism using parent report of narcissism between the U.S. and an Asian culture with a predominantly Chinese population (i.e., Hong Kong; Fung et al., 2010).

Currently, no study has yet to examine differences in self-report narcissism between children and adolescents in the U.S. and Singapore.

In regard to gender differences, studies have mostly found males to have higher levels of narcissism when compared to females, even among children and adolescents (Ang & Raine, 2009; Loke & Lowe, 2014). Despite this gender difference in overall narcissism, when comparisons across narcissistic dimensions in males and females were conducted, equivocal results have been reported (Fossati et al. 2010; Washburn et al., 2004). As such, this study seeks to provide clarity into the relationship between narcissism and gender.

Pertaining to age differences, only a few studies have looked into differences in narcissism during childhood and adolescence. For example, Foster et al.'s (2003) study found higher levels of narcissism in adolescents between 15 and 19 years as compared to children and adolescents between the ages of 8 and 14 years. However, a longitudinal study conducted by Carlson and Gjerde (2009) found that individuals had higher levels of narcissism when they were at age 18 as compared to when they were 14 years old. In addition, despite the relationship between adolescent egocentrism, which develops during adolescence, and narcissism (Aalsma et al., 2006; Lapsley et al., 1989), no study has examined differences in narcissism between preadolescents and early adolescents.

CHAPTER III

Method

The participants, measures, and procedures used in the study are described in detail in this chapter. In addition, data analyses that were performed to answer the research questions are discussed below.

Participants

Total Sample. This study was part of a larger study and consisted of archival data. Data were collected from a total of 717 participants (558 individuals from Singapore, and 159 individuals from the U.S.). Most participants (64.3%) were from the dominant ethnic group of each culture (i.e., White and Chinese). Students were between the ages of 12 and 14 years (M = 12.93, SD = 0.57*)*. Of the 717 students, 145 (20.2%) were 12 years old (preadolescents), and 572 (79.8%) were between the ages of 13 and 14 (early adolescents). There were a total of 386 (53.8%) males and 331 (46.2%) females in the study.

After data screening was completed, one missing datum and 15 outliers were removed from the dataset; this resulted in a total sample size of 701 participants (550 individuals from Singapore, and 151 individuals from the U.S.). A total of 451 participants (64.3%) were from the dominant ethnic group of each culture (i.e., White and Chinese). All students were between the ages of 12 and 14 years (M = 12.94, SD = .57). Of the 701 participants, 137 (19.5%) participants were 12 years old (preadolescents), and 564 (80.5%) were between the ages of 13 and 14 (early adolescents). There were 377 (53.8%) males and 324 (46.2%) females in the study.

Singapore Subsample. The Singapore subsample consisted of 558 participants (308 males and 250 females) between the ages of 12 and 14 (M = 13.06, SD = 0.49). The ethnic composition of this subsample included 60.2% Chinese, 26.5% Malay, 11.1% Indian, 0.2% Eurasian, and 2.0% other.

After data screening was completed, the Singapore subsample consisted of 550 participants (304 males and 246 females) between the ages of 12 and 14 (M = 13.06, SD = 0.49). The ethnic composition of this subsample included 60.5% Chinese, 26.4% Malay, 11.1% Indian, 0.2% Eurasian, and 1.8% other.

U.S. Subsample. The U.S. subsample consisted of 159 participants (78 males and 81 females) between the ages of 12 and 14 (M = 12.49, SD = 0.65). The ethnic composition of this subsample included 78.0% White, 3.8% African American, 2.5% Hispanic, 0.6% Native American, 0.6% Asian, and 14.5% other.

After data screening was completed, the U.S. subsample consisted of 151 participants (73 males and 78 females) between the ages of 12 and 14 (M = 12.51, SD = 0.65). The ethnic composition of this subsample included 78.1% White, 4.0% African American, 2.0% Hispanic, 0.7% Native American, 0.7% Asian, and 14.6% other.

Measures

NPQC-R. The NPQC-R is a 12-item, self-report measure of narcissism designed specifically for children and adolescents. The measure consists of a Superiority (6-items) subscale and an Exploitativeness (6-items) subscale. The Superiority subscale measures the individual's inflated sense of self, while the Exploitativeness subscale assesses the individual's tendency to manipulate and take advantage of others. A Total NPQC-R score, which measures total narcissism, is obtained by summing the raw scores on the Superiority and Exploitativeness subscales. All participants rated items on the NPQC-R using a 5-point Likert scale, ranging from 1 (*not at all like me)* to 5 (*completely like me)*.

Ang and Raine (2009) reported internal consistency reliability estimates of .84 to .85 for the Total NPQC-R scores, .73 to .77 for the Exploitativeness scores, and .79 for the Superiority

scores among Singapore students. In Loke and Lowe's (2014) study with U.S. students, the authors reported internal consistency reliability estimates of .85 for the Total NPQC-R scores, .85 for the Superiority scores, and .82 for the Exploitativeness scores.

Convergent and discriminant evidence of validity were also found with groups of Singapore and U.S. students. In Ang and Raine's (2009) study with Singapore students, the authors found small to medium effect sizes between total narcissism and exploitativeness scores, and aggression and proactive aggression scores; in contrast, they found negligible to small effect sizes between scores on total narcissism and its subscales (i.e., superiority and exploitativeness), and scores on teacher satisfaction, instrumental help, and teacher conflict. In Loke and Lowe's (2014) study with U.S. students, they found large effect sizes between scores on total narcissism, and scores on grandiosity and manipulativeness. Pertaining to the NPQC-R's subscales, moderate effect sizes between superiority scores and grandiosity scores, and a large effect size between exploitativeness scores and manipulativeness scores were also found. In contrast, negligible to small effect sizes between total narcissism, superiority, and exploitativeness scores, and trait anxiety scores were found.

Procedures

The data collection procedures for both the Singapore and U.S. subsamples are described below. The University of Kansas (KU) Human Subjects Committee – Lawrence campus (HSC-L) approval was obtained to analyze the archival data.

Singapore Subsample. Approval for this study was granted by the Nanyang Technological University's (NTU) Institutional Review Board (IRB), and permission to collect data was granted by Singapore's Ministry of Education. Principals from four schools (two schools from the Northern zone, and two schools from the Southern zone) in Singapore were

contacted. However, as this study was part of a larger study, data were collected only from two schools (one from the Northern zone, and one from the Southern zone) for the NPQC-R. The school counselor from the two schools who collected the NPQC-R data was then contacted to aid in selecting classes from each grade level that were representative of the student population in the school. A passive consent procedure was used to obtain participation from parents, and verbal assent was obtained from all participants prior to the administration of the questionnaires. Participation in this study was completely voluntary, and participants were explicitly informed that they could choose not to participate or withdraw their participation in this study at any time without penalty. Four test administrators were involved in the process of collecting data in this study, and followed standardized test administration procedures in the administration of the measure. Test administrators read scripted instructions as described on the NPQC-R prior to the administration of the measure. They were also present during the test administration process to answer questions the participants had.

U.S. Subsample. This study was part of a larger study that was approved by KU's HSC-L. Once approval was obtained from the HSC-L, permission to collect data was sought from either the school district's superintendent and/or the school principal in the Midwestern region of the U.S. Eight school principals agreed to allow students from their school to participate in this study. In order for students to participate in this study, written informed consent was obtained from the students' parents at least one week prior to the administration of measures. In addition, before the measures were administered, two test administrators obtained verbal assent from all participants. Participation in this study was completely voluntary, and students were informed that they could choose to not participate in this study, or could withdraw from this study at any point in time without penalty. Test administrators followed standardized test administration

procedures in the administration of the measures, and read instructions described on the NPQC-R prior to the administration of this measure. Both test administrators were also present during the test administration process to answer any questions the participants had.

Data Analyses

This section describes how data were assessed for missing values, outliers, and normality using SPSS statistical software (Version 22.0; IBM Corp., 2013). Next, procedures used to test for construct bias on the NPQC-R using multi-group CFAs and differences in internal consistency reliability estimates between groups are described. The analyses for outliers, normality, and differences in internal consistency reliability estimates were conducted using SPSS statistical software (Version 22.0), whereas the multi-group CFA analyses were conducted using Mplus, Version 7.11 (Muthén & Muthén, 1998-2013). In addition, Mplus, Version 7.11 was used to perform latent means analyses across groups of interests, i.e., culture (Singapore, U.S.), gender (males, females), and age (preadolescents, adolescents).

Data Screening. Descriptive analysis performed indicated that there was one missing datum in the data. In line with Allison's (2009) suggestion that the use of listwise deletion is the most efficient technique if missing data consists of less than 5% of the total sample, data from that participant were removed.

Next, the data were examined for outliers and normality. Leys, Ley, Klein, Bernard and Licata (2013) recommends the detection of outliers using the Median Absolute Deviation (MAD) method as they claim that this method is one of the most robust measures of dispersion. Analysis using the median plus or minus 2.5 times the MAD method resulted in 15 cases being deleted, and this yielded a total sample of 701 participants. To test for normality in the data, the Shapiro-Wilks test was performed. A significant test statistic was obtained for all items on the NPQC-R

(p <.001) and the NPQC-R subscales (Superiority, p <.001; Exploitativeness, p <.001). As such, these results indicate that the data are not normally distributed.

The mean- and variance-adjusted weighted least squares (WLSMV) estimator was used to estimate model parameters because it is an unbiased estimator for non-normal data (Flora & Curran, 2004). In addition, the WLSMV estimator was selected because of its robustness in dealing with ordinal data. Jamieson (2004) suggests that a Likert scale is in fact a set of ordered categories as intervals between values on a Likert scale cannot be assumed to be equal. As such, with the NPQC-R being measured on a 5-point Likert scale, the WLSMV estimator was used.

Tests for Construct Bias. To test for construct bias in the NPQC-R across culture, gender, and age, measurement invariance analyses between groups were performed. A CFA on the total sample was first performed to determine the model that best fits the data. Single-group CFAs were next conducted to obtain models acceptable for groups of interests before multi-group CFAs were performed (Brown, 2015). In addition, tests to examine differences in internal consistency reliability estimates between groups were performed.

Total Sample and Single-Group CFAs. A CFA on the total sample was first performed to determine the model that best fits the data. Thereafter, single-group CFAs based on the best-fitting model obtained from the previous analysis were performed across the different groups of interests. In other words, single-group CFAs were conducted separately for the Singapore subsample and the U.S. subsample to obtain a model that is acceptable for both groups. This process was then repeated for males and females, as well as for preadolescents and adolescents.

Models were then evaluated for goodness-of-fit using numerous methods. These included examining the chi-square (χ^2) statistic, the Comparative Fit Index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1993), which is also known as the Nonnormed Fit

Index (NNFI; Bentler & Bonett, 1980), and the Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993) and its 90% confidence interval (CI).

If a non-significant χ^2 estimate results when the CFA is performed, this would indicate an overall good model fit. However, χ^2 estimates are inflated by large sample sizes, such that χ^2 estimates are routinely rejected despite minimal differences between sample variances/covariances and model-implied variances/covariances (Brown, 2015). In this study, the total sample size was large (n = 701).

In regard to the other criteria to assess for model fit, some researchers have suggested that CFI and TLI values that are above .90 are indicative of acceptable model fit (Bentler, 1989; Bentler & Bonett, 1980), while CFI and TLI values that are close to .95 are indicative of good model fit (Hu & Bentler, 1999). Browne and Cudeck (1993) recommended that RMSEA values greater than .10 indicate poor model fit. In addition, RMSEA values between .08 and .10 are indicative of mediocre fit (MacCallum, Browne, & Sugawara, 1996), whereas RMSEA values between .05 and .08 are indicative of fair fit (Browne & Cudeck, 1993). Browne and Cudeck (1993) further suggested that RMSEA values less than .05 are indicative of close fit.

Multi-group CFAs. Multi-group CFAs were then conducted to test for invariance across culture, gender, and age. This entailed testing a series of nested models by comparing a less restrictive model with a more restricted model that has parameters increasingly constrained. The more restricted model is said to be nested within the less restricted model (Bollen, 1989). Using a stepwise procedure, tests for configural invariance, weak invariance, and strong invariance were performed in this order.

First, configural invariance was tested across groups. Configural invariance tests whether the same measurement model, i.e., the same number of factors and factor-indicators

correspondence, is similar across both groups (Brown, 2015). If the model fits the data, this would indicate that configural invariance is established across groups, and that the factor structure for both groups is similar.

Once configural invariance has been established, tests for weak invariance were performed. This entailed constraining the factor loadings on both groups to be equal. Should weak invariance across groups be established, this would indicate that factor loadings are proportionally equal between groups (Brown, 2015). Tests for strong invariance were next performed.

Strong factorial invariance imposes additional constraints on top of the weak factorial invariance model, such that indicator thresholds were constrained to be equal across groups. Should strong factorial invariance be established, this would indicate that the latent factors of both groups have the same unit of measurement (or are on the same scale).

There are numerous methods to evaluate whether the competing nested model fits the data and model. A χ^2 difference test has been traditionally used to compare nested models, with a non-significant χ^2 estimate ($p > .01$) demonstrating evidence of invariance between groups (Cheung & Rensvold, 2002). However, alternative fit criteria have been suggested as basing invariance testing on a χ^2 difference test may be too stringent for applied research, with even small differences between nested models resulting in a significant χ^2 estimate when large sample sizes are used (Cheung & Rensvold, 2002; King, Ganotice, & Watkins, 2012). Other alternative fit criteria include: a decrease in CFI of less than or equal to .01 (Cheung & Rensvold, 2002), a decrease in RMSEA of less than .015 (Chen, 2007) and a RMSEA value of the alternative model falling within the 90% CI of the null model (Ginevra et al., 2013; Little, 1997). Cheung and Rensvold (2002) recommended reporting the change in the CFI, ΔCFI, fit statistic as this statistic

is neither influenced by sample size or model complexity, while Little and colleagues (2007) recommended that the RMSEA fit criterion is most appropriate if the goal of a study is to identify an invariance model that best fits the data.

Tests for Internal Consistency Reliability between Groups. Another method to test for construct bias is to examine internal consistency reliability estimates between groups. Internal consistency reliability estimates for the NPQC-R Total score and its subtests were first obtained for each group of interest. Next, the Feldt technique was employed to test for differences in internal consistency reliability between the groups (Feldt, 1969). The Feldt technique provides an *F*-test statistic, which is derived by dividing the larger error variance by the smaller error variance for groups of interest on the NPQC-R Total and its subscales *(see equation 1)*.

$$F = \frac{1 - alpha_1}{1 - alpha_2} \tag{1}$$

This *F*-test statistic was then compared to the critical *F*-value from the *F*-distribution table. If the *F*-test statistic did not exceed the critical *F*-value, this would indicate that there were no significant differences between the two groups. The Feldt technique was performed on the NPQC-R Total scale and its subscales for the different groups of interest. If non-significant differences were found, then these findings would suggest that there is no construct bias across the groups. In addition, effect sizes for group differences were calculated. According to Cohen (1988)'s guidelines, effect size estimates of 0.2, 0.5, and 0.8 represent small, medium, and large effect sizes, respectively. As such, a negligible effect size, or an effect size estimate of less than

0.2, would indicate minimal to no differences between groups on the NPQC-R scale and subscales, and would suggest that there is no construct bias between groups.

Furthermore, the 95% CI for groups of interest on the NPQC-R Total and its subscales were examined. An overlap between the groups of interests on each scale and subscale would suggest that the internal consistency reliability estimates are similar across groups of interest on the NPQC-R Total and its subscales (Lowe, 2014).

Comparison of Latent Factors between Groups. Once strong factorial invariance has been established between groups of interest, comparisons of the latent means between groups may be conducted. The latent means of one of the comparison groups were constrained, while the latent means of the other group were not be restricted (Brown, 2015). The group with latent means constrained served as the reference group. Latent mean estimates for the group with latent means freely estimated represent deviations from the reference group. As such, positive latent mean estimates for the group with latent means freely estimated indicate higher mean scores for this group as compared to the reference group.

Summary

Data collected from 717 participants between the ages of 12 and 14 years in the U.S. and Singapore were screened for missing data, outliers, and normality using SPSS (Version 22.0). A total of 16 cases were removed from the dataset due to missing data and outliers, resulting in the analysis of data from 701 participants. All participants completed the NPQC-R as part of a larger study. Tests for construct bias using measurement invariance and difference in internal consistency reliability estimates across culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents) were conducted. Measurement invariance analyses were performed using Mplus, Version 7.11 (Muthén & Muthén, 1998-2013). After single-group CFAs

were performed, multi-group CFAs were conducted. Multi-group CFAs entailed testing a series of nested models in a stepwise manner, with tests for configural invariance, weak invariance, and strong invariance performed in this order. Differences in internal consistency reliability estimates between groups were tested using the Feldt technique (Feldt, 1969). In addition, the overlap in the 95% CI between the groups' internal consistency reliability estimates was examined. Last, comparisons of latent means between groups were conducted once measurement invariance was found to be tenable between different groups.

CHAPTER IV

Results

This chapter describes the results obtained from the analyses performed using the factor analysis and reliability methods to examine construct bias. According to the factor analysis method, tests of measurement invariance between culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents) were performed to determine whether construct bias was present between groups. In addition, according to the reliability method, internal consistency reliability estimates on the NPQC-R Total and subscales between groups were examined. Two internal consistency reliability methods were utilized to examine construct bias in this study. Specifically, overlaps in the 95% confidence intervals around the internal consistency reliability estimates were examined between groups, and tests of internal consistency reliability estimates between groups using the Feldt technique were performed. Once at least partial measurement invariance was obtained, latent means analyses between groups were performed to determine whether differences between groups on the latent factors were present. Data analyses presented in this section are organized according to the research questions proposed.

Research Question 1

The first research question of this study seeks to examine whether there is construct bias in the NPQC-R scores between groups using the factor analysis method. Specifically, tests of measurement invariance between groups were performed in this study.

Prior to performing tests of measurement invariance, confirmatory factor analyses (CFA) were performed. CFAs were performed for the total sample and the subsamples (i.e., Singapore, U.S., males, females, preadolescents, early adolescents) using the WLSMV estimator to estimate model parameters. The preponderance of evidence based on the following criteria were used to

determine acceptable model fit: 1) a non-significant χ^2 estimate (Brown, 2015), 2) CFI and TLI or NNFI values that are at least above .90 (Bentler, 1989; Bentler & Bonett, 1980), and 3) RMSEA values that are less than or equal to .10 (MacCallum et al., 1996). However, as mentioned before, χ^2 estimates tend to be de-emphasized as significant χ^2 estimates are routinely obtained when the total sample size is large (Brown, 2015).

A CFA was performed on the total sample to determine the model that best fits the data. The unmodified one-factor model had a poor fit to the data, (WLSMVχ^2 (54, n = 701) = 676.63, p < .001, CFI = .87, TLI/NNFI = .85, RMSEA = .13). Standardized coefficients ranged from .36 to .75. Based on Brown's (2015) recommendation of a "salient" factor loading being defined by standardized coefficients that are .30 and above, factor loadings in this model were adequate (see Table 1).

Next, a CFA was performed on the unmodified two-factor model, which was validated with a sample of Singapore students. Results from this unmodified two-factor model had improved model fit, (WLSMVχ^2 (53, n = 701) = 241.92, p < .001, CFI = .96, TLI/NNFI = .95, RMSEA = .07). Standardized coefficients ranged from .42 to .81, and this suggests that the factor loadings were adequate (Brown, 2015; see Table 1). However, the unmodified two-factor model had one large modification index value, i.e., 76.70. Further inspection of the large modification index value indicated a factor cross-loading (i.e., an item loaded on more than one factor). This suggests that in addition to its loading on the Exploitativeness factor, Item 2 on the NPQC-R also loads on the Superiority factor. Item 2 states "If I ruled the world, it would be a better place". This finding is similar to Loke and Lowe's (2014) results obtained from a CFA conducted on the NPQC-R scores among U.S. students. Furthermore, Ang and Yusof (2006) indicated that it was plausible that the item content of Item 2 may tap on to students' beliefs about their own

Table 1

Standardized Factor Loadings of the 1-factor (Unmodified), 2-factor (Unmodified), and 2-factor (Modified) Models

NPQC-R subscales and items	1-factor	2-factor (unmodified model)	2 factor (modified) model
Superiority			
1. I always know what I am doing.	.38	.42	.42
3. I am going to be a great person.	.64	.68	.69
6. I was born a good leader.	.67	.71	.71
8. I am really a special person.	.73	.77	.77
10. I think my body looks good.	.59	.63	.63
11. I think I am a great person.	.75	.79	.79
2. If I ruled the world it would be a better place.			.34
Exploitativeness			
2. If I ruled the world it would be a better place.	.49	.54	.22
4. I am good at getting people to do things my way.	.72	.81	.83
5. It is easy for me to control other people.	.63	.75	.76
7. I would do almost anything if you dared me.	.38	.44	.45
9. I can make people believe anything I want them to.	.60	.67	.68
12. When I am supposed to be punished, I can usually talk my way out of it.	.36	.43	.44

superiority. In Ang and Yusof's (2006) principal component analysis of Singapore students, Item 2 had a high secondary loading on the Superiority factor. However, Byrne (1994) cautioned that model respecification should only be made if respecification made statistical and substantive sense. In this study, model respecification made statistical sense due to the large modification index value found. It also made substantive sense based on previous findings of Item 2 having a high secondary loading on the Superiority factor in the Singapore sample (Ang & Yusof, 2006) and evidence of a factor cross-loading in the CFA performed in the U.S. sample (Loke & Lowe, 2014). As such, a CFA was performed on the modified two-factor model, and results indicated that the fit indices for this model had the best model fit to the data in the present study (WLSMVχ^2 (52, n = 701) = 168.56, p < .001, CFI = .98, TLI/NNFI = .97, RMSEA = .06). In addition, standardized coefficients were largely adequate, with the factor cross-loading coefficient for Item 2 on the Exploitativeness subscale being .34 (see Table 1).

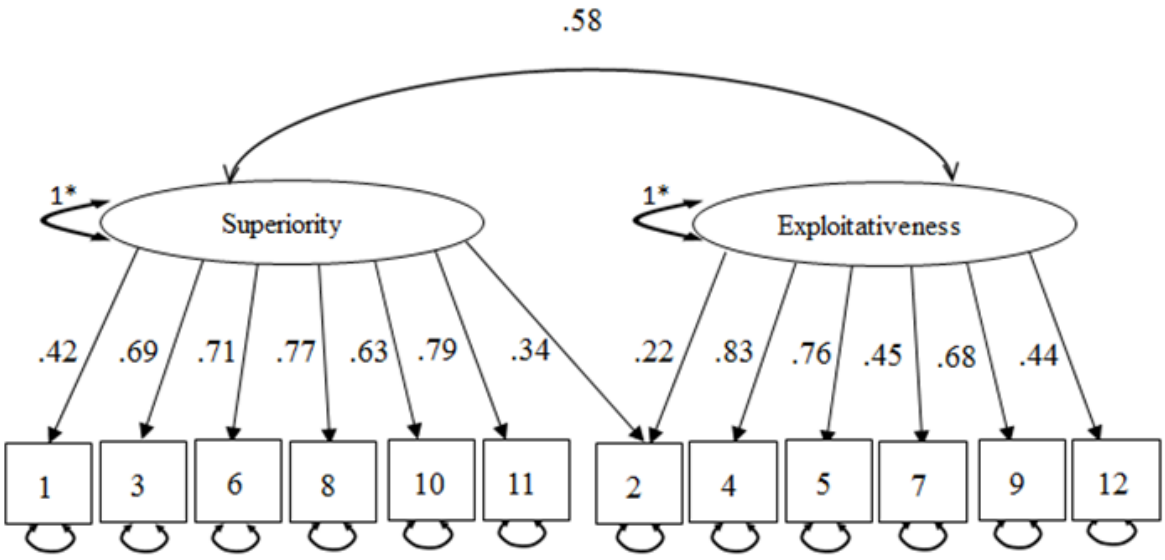

Figure 1. Modified 2-factor (superiority and exploitativeness) confirmatory factor measurement model for the 12-item NPQC-R measure

Taken together, the model fit for the modified two-factor model with one indicator cross-loading was adequate. Standardized factor coefficients are presented for the total sample in Table 1 and model fit statistics for the total sample are shown in Table 2. Figure 1 shows the modified two-factor model.

Table 2

Fit Indices for the Confirmatory Factor Analyses of the Narcissistic Personality Questionnaire for Children - Revised (NPQC-R) for the Total Sample (n=701)

Model	WLSMV χ^2	df	CFI	TLI/NNFI	RMSEA [90% CI]
1 factor (unmodified)	676.63**	54	.87	.85	.13 [.12, .14]
2 factor (unmodified)	241.92**	53	.96	.95	.07 [.06, .08]
2 factor (modified)	168.56**	52	.98	.97	.06 [.05, .07]

Note. WLSMVχ^2 = Mean- and Variance-adjusted Weighted Least Squares Chi-square; CFI =Comparative Fit Index; TLI/NNFI = Tucker-Lewis Index/Nonnormed Fit Index; RMSEA = Root Mean Square Error of Approximation; *$p < .05$, ** $p < .01$

The modified two-factor model was examined with all groups (i.e. Singapore, U.S., males, females, preadolescents, early adolescents) prior to performing measurement invariance. Fit indices for the modified two-factor model for the subsamples are presented in Table 3. Acceptable model fit for the modified two factor model was found for males (WLSMVχ^2 (52, n = 376) = 167.74, $p < .001$, CFI = .95, TLI/NNFI = .93, RMSEA = .08), females (WLSMVχ^2 (52, n = 324) = 77.18, $p = .01$, CFI = .99, TLI/NNFI = .99, RMSEA = .04), Singapore students (WLSMVχ^2 (52, n = 550) = 184.52, $p < .001$, CFI = .96, TLI/NNFI = .95, RMSEA = .07), U.S.

students (WLSMVχ^2 (52, n = 151) = 91.59, p <.001, CFI = .97, TLI/NNFI = .96, RMSEA = .07), preadolescents (WLSMVχ^2 (52, n = 137) = 95.52, p < .001, CFI = .97, TLI/NNFI = .97, RMSEA = .08), and early adolescents (WLSMVχ^2 (52, n = 564) = 140.95, p < .001, CFI = .98, TLI/NNFI = .97, RMSEA = .06).

Table 3

Fit Indices for the Modified Two-Factor Models of the Narcissistic Personality Questionnaire for Children - Revised (NPQC-R)

Groups	WLSMV χ^2	Df	CFI	TLI/NNFI	RMSEA [90% CI]
Singapore (n=550)	184.52**	52	.96	.95	.07 [.06, .08]
U.S. (n=151)	91.59**	52	.97	.96	.07 [.05, .10]
Males (n=376)	167.74**	52	.95	.93	.08 [.06, .09]
Females (n=324)	77.18*	52	.99	.99	.04 [.02, .06]
Preadolescent (n=137)	95.52**	52	.97	.97	.08 [.05, .10]
Early Adolescent (n=564)	140.95**	52	.98	.97	.06 [.04, .07]
U.S. Males (n=73)	94.19**	52	.95	.93	.11 [.07, .14]
U.S. Females (n=78)	73.37*	52	.97	.96	.07 [.03, .11]
Singapore Males (n=304)	175.11**	52	.92	.90	.09 [.07, .10]
Singapore Females (n=246)	91.39**	52	.98	.97	.06 [.04, .07]

Note. WLSMVχ^2 = Mean- and Variance-adjusted Weighted Least Squares Chi-square; CFI = Comparative Fit Index; TLI/NNFI = Tucker-Lewis Index/Nonnormed Fit Index; RMSEA = Root Mean Square Error of Approximation; *p < .05, ** p < .01

As acceptable model fit was obtained for all subsamples, measurement invariance testing across culture, gender, and age were performed. Measurement invariance is an important concept in test bias (Reynolds & Keith, 2013). If the test is invariant across groups, this suggests that the test is unbiased, in that individuals from different groups with similar latent traits should have similar observed scores (Meredith, 1993; Reynolds & Keith, 2013). Only when measurement invariance has been satisfied can groups be compared, as this suggests that the comparison groups have similar form, factor loadings, and thresholds (Brown, 2015). Measurement invariance exists when the preponderance of goodness-of-fit criteria are met: 1) a non-significant ΔWLSMV χ^2, 2) a decrease in CFI of less than or equal to .01 (Cheung & Rensvold, 2002), 3) a decrease in RMSEA of less than .015 (Chen, 2007), and a RMSEA value of the alternative model falling within the 90% CI of the null model (Ginevra et al., 2013; Little, 1997). Table 4 presents the results of tests of invariance for form (i.e., configural invariance), form and factor loadings (i.e., weak invariance), and form, factor loadings and thresholds (i.e., strong invariance). In addition, Table 5 presents the means and standard deviations for the NPQC-R scores for the groups.

Research Question 1a. Research question 1a seeks to examine whether construct bias is present across culture using the factor analysis method. To address this research question, tests of measurement invariance across Singapore and U.S. subsamples were performed. Configural invariance between Singapore and U.S. students was demonstrated (WLSMVχ^2 (104, n = 701) = 261.50, p < .001, CFI = .97, RMSEA = .07), and suggests invariance of form between the two groups. Next, test of weak invariance was performed between the two groups. Results found invariance of form and factor loadings between Singapore and U.S. students (ΔWLSMV χ^2 (11) = 45.88, p < .001; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null

Table 4

Testing for Measurement Invariance: Results of Multigroup Confirmatory Factor Analysis (n = 701)

Model	WLSMV χ^2	df	CFI	RMSEA [90% CI]
Cultural invariance (Singapore, U.S.)				
Model 1: Model with no equality constraints imposed	261.50**	104	.97	.07 [.06, .08]
Model 2: Model with factor loadings constrained	303.43**	115	.96	.07 [.06, .08]
Model 3: Model 2 with factor loading and thresholds constrained	477.65**	161	.93	.08 [.07, .08]
Model 4: Model 3 with item 1 threshold 4 freed	444.86**	160	.94	.07 [.06, .08]
Model 5: Model 4 with item 1 threshold 3 freed	420.61**	159	.94	.07 [.06, .08]
Model 6: Model 5 with item 10 thresholds 3 and 4 freed	407.22**	157	.95	.07 [.06, .08]
Model 7: Model 6 with item 12 threshold 1 freed	394.44**	156	.95	.07 [.06, .07]

Model 2 - Model 1 (ΔWLSMV χ^2 (11) = 45.88, $p < .01$; ΔCFI $\leq .01$; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)

Model 3 - Model 2 (ΔWLSMV χ^2 (46) = 200.28, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .08 within 90% CI of null model)

Model 4 - Model 2 (ΔWLSMV χ^2 (45) = 164.79, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)

Model 5 - Model 2 (ΔWLSMV χ^2 (44) = 137.29, $p < .01$; ΔCFI > .01[a]; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)

Model 6 - Model 2 (ΔWLSMV χ^2 (42) = 122.22, $p < .01$; ΔCFI > .01[a]; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)

Table 4 (cont'd)

Testing for Measurement Invariance: Results of Multigroup Confirmatory Factor Analysis (n = 701)

Model	WLSMV χ^2	df	CFI	RMSEA [90% CI]
Model 7 - Model 2 (ΔWLSMV χ^2 (41) = 108.44, p < .01; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Gender invariance (Male, Female)				
Model 1: Model with no equality constraints imposed	250.14**	104	.97	.06[.05, .07]
Model 2: Model with factor loadings constrained	235.96**	115	.98	.06 [.05, .07]
Model 3: Model 2 with factor loading and thresholds constrained	336.36**	161	.97	.06 [.05, .06]
Model 4: Model 3 with item 2 threshold 4 freed	323.16**	160	.97	.05 [.05, .06]
Model 2 - Model 1 (ΔWLSMV χ^2 (11) = 14.00, p = .23; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .06 within 90% CI of null model)				
Model 3 - Model 2 (ΔWLSMV χ^2 (46) = 109.61, p < .01; ΔCFI > .01[a]; ΔRMSEA < .015; RMSEA of .06 within 90% CI of null model)				
Model 4 - Model 2 (ΔWLSMV χ^2 (45) = 94.99, p < .01; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .05 within 90% CI of null model)				
Age invariance (Late Childhood, Early Adolescence)				
Model 1: Model with no equality constraints imposed	231.94**	104	.98	.06 [.05, .07]
Model 2: Model with factor loadings constrained	240.11**	115	.98	.06 [.05, .07]
Model 3: Model 2 with factor loading and thresholds constrained	308.70**	161	.97	.05 [.04, .06]

Table 4 (cont'd)

Testing for Measurement Invariance: Results of Multigroup Confirmatory Factor Analysis (n = 701)

Model	WLSMV χ^2	df	CFI	RMSEA [90% CI]
Model 2 - Model 1 (ΔWLSMV χ^2 (11) = 24.14, p = .01; ΔCFI $\leq$ .01; ΔRMSEA < .015; RMSEA of .06 within 90% CI of null model)				
Model 3 - Model 2 (ΔWLSMV χ^2 (46) = 79.27, p < .01; ΔCFI $\leq$ .01; ΔRMSEA < .015; RMSEA of .05 within 90% CI of null model)				
Culture and Gender invariance (Singapore male, U.S. male, Singapore female, U.S. female)				
Model 1: Model with no equality constraints imposed	418.52**	208	.96	.08 [.07, .09]
Model 2: Model with factor loadings constrained	460.07**	241	.96	.07 [.06, .08]
Model 3: Model 2 with factor loading and thresholds constrained	760.35**	379	.92	.08 [.07, .08]
Model 4: Model 3 with item 2 threshold 4 freed	737.75**	376	.93	.07 [.07, .08]
Model 5: Model 4 with item 1 threshold 4 freed	708.38**	373	.93	.07 [.06, .08]
Model 6: Model 5 with item 12 threshold 1 freed	691.01**	370	.93	.07 [.06, .08]
Model 7: Model 6 with item 2 threshold 3 freed	678.39**	367	.94	.07 [.06, .08]
Model 8: Model 7 with item 10 thresholds 3 and 4 freed	656.07**	361	.94	.07 [.07, .08]
Model 9: Model 8 with item 1 threshold 3 freed	635.06**	358	.94	.07 [.06, .08]
Model 10: Model 9 with item 12 threshold 4 freed	627.76**	355	.94	.07 [06, .08]

Table 4 (cont'd)

Testing for Measurement Invariance: Results of Multigroup Confirmatory Factor Analysis (n = 701)

Model	WLSMV χ^2	df	CFI	RMSEA [90% CI]
Model 11: Model 10 with item 12 threshold 3 freed	619.85**	352	.95	.07 [.06, .07]
Model 2 - Model 1 (ΔWLSMV χ^2 (33) = 70.27, p < .01; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 3 - Model 2 (ΔWLSMV χ^2 (138) = 357.05, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .08 within 90% CI of null model)				
Model 4 - Model 2 (ΔWLSMV χ^2 (135) = 330.71, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 5 - Model 2 (ΔWLSMV χ^2 (132) = 293.96, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 6 - Model 2 (ΔWLSMV χ^2 (129) = 272.70, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 7 - Model 2 (ΔWLSMV χ^2 (126) = 257.82, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 8 - Model 2 (ΔWLSMV χ^2 (120) = 230.20, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 9 - Model 2 (ΔWLSMV χ^2 (117) = 203.58, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 10 - Model 2 (ΔWLSMV χ^2 (114) = 194.68, p < .01; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				
Model 11 - Model 2 (ΔWLSMV χ^2 (111) = 184.84, p < .01; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model)				

Note. WLSMV χ^2 = Mean- and Variance-adjusted Weighted Least Squares χ^2; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; [a]all values are rounded to 2 decimal points but ΔCFI was greater than .01; **$p < .01$

Table 5

Means (*M*) *and Standard Deviations (SD) of the Narcissistic Personality Questionnaire for Children-Revised (NPQC-R) Scores by Group*

	Superiority		Exploitativeness		Total NPQC-R	
	M	*SD*	*M*	*SD*	*M*	*SD*
Total Sample (*n*=701)	16.64	5.27	12.56	4.28	29.20	8.24
Singapore (*n* = 550)	15.65	4.87	12.37	4.28	28.02	7.99
U.S. (*n* = 151)	20.26	5.09	13.23	4.24	33.50	7.73
Males (*n* = 377)	16.79	5.12	12.88	4.47	29.67	8.15
Females (*n* = 324)	16.48	5.44	12.18	4.03	28.66	8.33
Preadolescents (*n* = 137)	18.07	5.77	12.53	4.44	30.60	8.77
Early Adolescents (*n* = 564)	16.30	5.08	12.57	4.25	28.86	8.08
Singapore Males (*n* = 304)	15.92	4.72	12.85	4.44	28.77	7.97
Singapore Females (*n* = 246)	15.31	5.04	11.79	4.01	27.10	7.93
U.S. Males (*n* = 73)	20.37	5.20	13.04	4.64	33.41	7.88
U.S. Females (*n* = 78)	20.17	5.01	13.41	3.86	33.58	7.63

Note. Range of scores on the subscales is 6 to 30, and the full scale is 12 to 60. Higher mean scores indicate higher levels of narcissism.

model). Due to invariance of form and factor loadings between the two groups, test of strong invariance was next performed. Results did not suggest strong invariance (ΔWLSMV χ^2 (46) = 200.28, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .08 within 90% CI of the null model), based on Chen's (2007) assertion that the ΔCFI should be used as the main criterion in tests of

measurement invariance as it is less influenced by sample size and model complexity. Visual inspection of the modification indices were examined to identify if there were any large modification indices (M.I. > 3.84) present (Brown, 2015). Inspection of the modification indices indicated that threshold 4 of Item 1 (i.e., I always know what I am doing) was not equivalent between the groups, and suggests that a greater level of latent response is required before U.S. students will endorse the option of "completely like me". After threshold 4 of Item 1 was released, results did not suggest strong measurement invariance between the two groups (ΔWLSMV χ^2 (45) = 164.79, $p < .001$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). A subsequent visual inspection of the modification indices indicated that threshold 3 of Item 1 was not equivalent between the groups, and suggests that a greater level of latent response is required before U.S. students will endorse the option of "very much like me". As such, threshold 3 of Item 1 was also relaxed, but results still did not indicate strong invariance between Singapore and U.S. students (ΔWLSMV χ^2 (44) = 137.29, $p < .001$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). As strong invariance was not obtained, another visual inspection of the modification indices was conducted. Examination of the modification indices indicated that threshold 3 of Item 10 (i.e., My body looks good) was not equivalent between Singapore and U.S. students. Hoffman (2014) recommended that should the threshold to be released not be at either endpoint, that threshold and the threshold closest to the extreme end should be released because the middle threshold and the threshold closest to the endpoint are interdependent. As such, thresholds 3 and 4 of Item 10 were released simultaneously. This suggests that a greater level of latent response is required before Singapore students will endorse the options of "very much like me" and "completely like me" when compared to U.S. students. After releasing thresholds 3 and 4 of Item 10, results still did not

indicate strong invariance between Singapore and U.S. students (ΔWLSMV χ^2 (42) = 122.22, p < .001; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). As strong invariance between the two groups was still not obtained, visual inspection of modification indices was conducted once more. Examination of modification indices indicated that threshold 1 of Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) was not equivalent between groups, and this difference suggests that U.S. students are more likely to endorse the option of "not at all like me" as compared to Singapore students. As such, in addition to relaxing the constraints on thresholds 3 and 4 of Item 1 and Item 10, the constraint on threshold 1 of Item 12, was also released. Thereafter, the model with form and factor loadings constrained was compared to the model with form, factor loadings, and thresholds, with the exception of the thresholds for the respective items that were relaxed, constrained. Results supported findings of partial strong invariance on the NPQC-R scores between the U.S. and Singapore groups (ΔWLSMV χ^2 (41) = 108.44, p < .001; ΔCFI $\leq$ .01; ΔRMSEA < .015; RMSEA of .07 falling within 90% CI of null model). Partial measurement invariance is obtained when most parameters in the model are constrained to be invariant, while the remaining parameters are allowed to vary between groups (Byrne, Shavelson, & Muthén, 1989).

Research Question 1b. Research question 1b seeks to examine whether construct bias was present across gender using the factor analysis method. To address this research question, tests of measurement invariance across males and females were performed.

Configural invariance was found between males and females on the NPQC-R (WLSMVχ^2 (104, *n* = 701) = 250.14, *p* < .001, CFI = .97, RMSEA = .06), and this suggests invariance of form between the two groups. Next, test of weak invariance was performed. Results indicated an invariance of form and factor loadings between males and females

(ΔWLSMVχ^2 (11) = 14.00, p = .23; ΔCFI ≤ .01, ΔRMSEA < .015, and the alternative model's RMSEA of .06 falling within the 90% CI of the null model). As weak invariance was obtained, invariance of form, factor loadings, and thresholds was tested. Results from this test did not suggest that there was strong invariance between the two groups (WLSMVχ^2 (46) = 109.61, p < .01; ΔCFI > .01, ΔRMSEA < .015, and the alternative model's RMSEA of .06 falling within the 90% CI of the null model). Visual inspection of the modification indices was next undertaken to identify whether large modification indices (M.I. > 3.84) were present (Brown, 2015). Inspection of the modification indices indicated that threshold 4 associated with Item 2 (i.e., If I ruled the world, it would be a better place) was not equivalent for males and females. In this instance, a greater level of latent response is required before females will endorse the option of "completely like me". As such, the constraint on this threshold was removed. Results comparing the model with form, factor loadings, and thresholds, except for threshold 4 of Item 2, constrained, with the model with form and factor loadings constrained suggested that partial strong invariance was supported between males and females (ΔWLSMV χ^2 (45) = 94.99, p < .001; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .05 falling within 90% CI of null model).

Research Question 1c. Research question 1c seeks to examine whether construct bias was present across age using the factor analysis method. To address this research question, tests of measurement invariance across preadolescents and early adolescents were performed.

Configural invariance was found between preadolescents and early adolescents on the NPQC-R (WLSMVχ^2 (104, n = 701) = 231.94, p < .001, CFI = .98, RMSEA = .06), and this suggests invariance of form between the two groups. Next, test of weak invariance was performed. Results supported invariance of form and factor loadings (ΔWLSMVχ^2 (11) = 24.14, p = .01; ΔCFI ≤ .01; ΔRMSEA < .015; RMSEA of .06 falling within 90% CI of null model). As

weak invariance was obtained, test of strong invariance was then performed. Results indicated that invariance of form, factor loadings, and thresholds was found (ΔWLSMV χ^2 (46) = 79.27, p < .01; ΔCFI $\leq$ .01; ΔRMSEA < .015; RMSEA of .05 falling within 90% CI of null model) between the two groups. In other words, strong invariance between preadolescents and early adolescents was tenable.

Other Tests of Measurement Invariance. A related pilot test was conducted to determine whether the narcissism construct was equivalent across the four subgroups (i.e., Singapore males, U.S. males, Singapore females, U.S. females). Tests of measurement invariance were performed on the NPQC-R across all four subgroups simultaneously. The finding of measurement invariance across the four subgroups would support the pooling of data from the groups in the previous analyses. For example, when testing for measurement invariance across gender in the previous analyses, the male group consisted of both Singapore and U.S. males, whereas the female group included both Singapore and U.S. females. However, it has to be noted that the results from this pilot test should be interpreted with caution. Although the total sample size was large (n = 701), the sample size for some of the subgroups were small. Indeed, the U.S. male subsample and the U.S. female subsample in this study included only 73 and 78 participants, respectively.

First, individual CFAs were performed for all groups (i.e., Singapore males, U.S. males, Singapore females, U.S. females) as acceptable model fit for each group has to be established prior to performing tests of measurement invariance (Brown, 2015). When CFAs were performed for males and females in the Singapore and U.S. samples, acceptable model fit were obtained for Singapore males (WLSMVχ^2 (52, n = 304) = 175.11, p < .001, CFI = .92, TLI/NNFI = .90, RMSEA = .09), Singapore females (WLSMVχ^2 (52, n = 246) = 91.39, p <.001, CFI = .98,

TLI/NNFI = .97, RMSEA = .06), and U.S. females (WLSMVχ^2 (52, n = 78) = 73.37, p = .03, CFI = .97, TLI/NNFI = .96, RMSEA = .07). For the U.S. males, results from the CFA performed, (WLSMVχ^2 (52, n = 73) = 94.19, p <.001, CFI = .95, TLI/NNFI = .93, RMSEA = .11, indicated that the RMSEA value for this subsample was beyond MacCallum et al.'s (1996) criteria for acceptable model fit (see Table 3). However, researchers have asserted that the RMSEA values are often elevated when small sample sizes are used, such that "true" models are often over-rejected and result in the false identification of poor model fit (Iacobucci, 2010; Kenny, Kaniskan, & McCoach, 2014). Owing to the small sample size in this subsample, the preponderance of evidence gathered from multiple goodness of fit indexes (i.e., CFI > .90 and TLI/NNFI > .90) suggest an acceptable model fit for the U.S. males subsample.

Thereafter, tests for measurement invariance were performed across all four groups simultaneously. The results of these tests are shown in Table 4. Results from the test of configural invariance demonstrated acceptable fit, (WLSMVχ^2 (208, n = 701) = 418.52, p < .001, CFI = .96, RMSEA = .08). Next, test of weak invariance was performed between groups, and results indicated an invariance of form and factor loadings between the two groups (ΔWLSMVχ^2 (33) = 70.27, p < .01; ΔCFI $\leq$ .01, ΔRMSEA < .015, and the alternative model's RMSEA of .07 fell within the 90% CI of the null model). As weak invariance was obtained, invariance of form, factor loadings and thresholds was tested. Results from this test did not suggest that there was strong invariance between the four groups (ΔWLSMVχ^2 (138) = 357.05, p < .01; ΔCFI > .01, ΔRMSEA < .015, and the alternative model's RMSEA of .08 fell within the 90% CI of the null model).

Visual inspection of the modification indices indicated that threshold 4 of Item 2 (i.e., If I ruled the world, it would be a better place) was not equivalent across groups, as Singapore males

were more likely to endorse the option of "completely like me". As such, the constraint on this threshold was removed. After relaxing this constraint, results still did not indicate strong invariance across groups (ΔWLSMV χ^2 (135) = 330.71, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). A subsequent visual inspection of the thresholds indicated that threshold 4 of Item 1 (i.e., I always know what I am doing) was not equivalent across groups, as Singapore males and females were more likely to endorse the option of "completely like me". After this additional constraint was relaxed, strong invariance across groups was still not tenable (ΔWLSMV χ^2 (132) = 293.96, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). Further inspection of the thresholds indicated that threshold 1 of Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) was not equivalent across groups, as Singapore females were less likely to endorse the option of "not at all like me" as compared to the other groups. However, after relaxing this constraint, strong invariance across groups was not tenable (ΔWLSMV χ^2 (129) = 272.70, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). An inspection of the thresholds indicated that threshold 3 on Item 2 (i.e., If I ruled the world, it would be a better place) was not equivalent across groups, as a greater level of latent response is required before the Singapore female group will endorse the option of "very much like me" as compared to the other groups. After relaxing this constraint, strong invariance across groups was still not tenable (ΔWLSMV χ^2 (126) = 257.82, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). Further examination of the thresholds indicated that threshold 3 of Item 10 (i.e., My body looks good) was not equivalent across groups, with U.S. males more likely to endorse the option of "very much like me". Based on Hoffman's (2014) recommendation that the extreme threshold closest to the threshold to be released should also be

freed due to interdependency between the middle threshold and the extreme threshold, thresholds 3 and 4 of Item 10 were released simultaneously. After these additional thresholds were released, strong invariance across groups was still not tenable (ΔWLSMV χ^2 (120) = 230.20, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). A subsequent examination of the thresholds indicated that threshold 3 of Item 1 (i.e., I always know what I am doing) was not equivalent across groups, with a greater level of latent response required before all groups except for the Singapore male and female groups will endorse the option of "very much like me". After this additional constraint was relaxed, strong invariance across groups was still not tenable (ΔWLSMV χ^2 (117) = 203.58, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). Thereafter, an examination of thresholds indicated that threshold 4 of Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) was not equivalent across groups. In this instance, the Singapore male group was more likely to endorse the option of "completely like me" as compared to the other groups. After relaxing this constraint, strong invariance was still not tenable across groups (ΔWLSMV χ^2 (114) = 194.68, $p < .01$; ΔCFI > .01; ΔRMSEA < .015; RMSEA of .07 within 90% CI of null model). A subsequent inspection of the thresholds indicated that threshold 3 of Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) was not equivalent across groups. This suggests the Singapore male group was also more likely to endorse the "very much like me" option as compared to the other groups. After this additional constraint was released, partial strong invariance across groups was tenable (ΔWLSMVχ^2 (111) = 184.84, $p < .01$; ΔCFI $\leq$ .01, ΔRMSEA < .015, and the alternative model's RMSEA of .07 falling within the 90% CI of the null model). Although partial strong measurement invariance was tenable across Singapore males, U.S. males, Singapore females, and U.S. females, findings from this pilot test should be

interpreted with caution due to the small sample sizes in the U.S. male and female subsamples, and future replication of these analyses is recommended. Results from these analyses are presented in Table 4 and the means and standard deviations for these four groups (i.e., Singapore males, U.S. males, Singapore females, U.S. females) are presented in Table 5.

Research Question 2

The second research question of this study seeks to examine whether there is construct bias in the NPQC-R scores between groups using the reliability methods. To address research question 2, internal consistency reliability estimates and the 95% confidence interval around each reliability estimate for each NPQC-R scale and subscale were computed for all groups (i.e., Singapore, U.S., males, females, preadolescents, early adolescents). Table 6 shows the coefficient alphas for each group and the 95% confidence intervals for the NPQC-R scale and subscale scores. In addition, the Feldt technique was employed to test for construct bias. Table 7 displays group comparisons using the Feldt technique on the scale and subscale scores of the NPQC-R.

Research Question 2a. Research question 2a seeks to examine whether construct bias was present across culture using the internal consistency reliability methods. First, the 95% confidence interval overlap in the internal consistency reliability estimates between the Singapore and U.S. groups on the scale and subscales of the NPQC-R were examined. The results indicated that there was an overlap in the 95% confidence intervals on the Superiority and Exploitativeness subscales, and on the NPQC-R Total scale. Second, tests of internal consistency reliability estimates using the Feldt technique also indicated that comparisons on the Exploitativeness subscale, F (150, 549) = 1.03, p = .40, d = .09 and the NPQC-R Total scale, F (150, 549) = 1.11, p = .20, d = .10, between the Singapore and U.S. groups did not indicate the

Table 6

Internal Consistency Reliability Estimates and their 95% Confidence Interval for the Narcissistic Personality Questionnaire for Children - Revised Scale and Subscale Scores by Group (n = 701)

Groups	NPQC-R Total	Superiority	Exploitativeness
Culture			
Singapore	.81 [.78, .83]	.76 [.73, .79]	.68 [.64, .72]
U.S.	.79 [.74, .84]	.81 [.75, .85]	.67 [.58, .75]
Gender			
Males	.79 [.76, .82]	.77 [.73, .80]	.67 [.62, .72]
Females	.84 [.81, .86]	.82 [.79, .85]	.69 [.63, .74]
Age			
Preadolescents	.83 [.79, .87]	.84 [.79, .88]	.71 [.63, .78]
Early Adolescents	.81 [.78, .83]	.78 [.75, .81]	.67 [.63, .71]

presence of bias. However, on the Superiority subscale, significant results using the Feldt technique suggest the presence of bias between the two groups, F (549, 150) = 1.26, p = .04, d = .10, but the effect size was negligible. A negligible effect size indicates that the significant difference reported on the NPQC-R Superiority subscale was not clinically meaningful, and this suggests that there is no construct bias on this subscale.

Research Question 2b. Research question 2b seeks to examine whether construct bias was present across gender using the internal consistency reliability methods. To address this research question, the 95% confidence interval overlap in the internal consistency reliability

Table 7

Group comparisons using the Feldt Technique on the Narcissistic Personality Questionnaire for Children-Revised scale and subscale scores (n = 701)

Group Comparisons	NPQC-R Total			Superiority			Exploitativeness		
	F-value	*p*-value	E.S.	*F*-value	*p*-value	E.S.	*F*-value	*p*-value	E.S.
Singapore vs U.S.	1.11	.20	0.10	1.26	.04	0.10	1.03	.40	0.09
Males vs Females	1.31	< .01	0.09	1.28	.01	0.09	1.06	.29	0.08
Late Childhood vs Early Adolescence	1.12	.21	0.10	1.38	.01	0.11	1.14	.18	0.10

Note. E.S. = Effect Size

estimates between males and females on each NPQC-R scale and subscale were examined. An overlap in the 95% confidence interval using the internal consistency reliability estimate method between males and females on the Superiority and Exploitativeness subscales, and on the NPQC-R Total scale, was found. In addition, a test of internal consistency reliability estimates using the Feldt technique found no bias on the Exploitativeness subscale, F (375, 323) = 1.06, p = .29, d = .08. On the Superiority subscale, F (375, 323) = 1.28, p = .01, d = .09, and on the NPQC-R Total scale, F (375, 323) = 1.31, p < .01, d = .09 significant results were reported. However, as the effect sizes were negligible, results did not support the presence of construct bias between the two groups on the Superiority subscale and the NPQC-R Total scale. Taken together, there was no construct bias between males and females on the Exploitativeness and Superiority subscales, and on the NPQC-R Total scale.

Research Question 2c. Research question 2c seeks to examine whether construct bias was present across age using the internal consistency reliability methods. To address this research question, the 95% confidence interval overlap in the internal consistency reliability estimates were investigated between preadolescents and early adolescents. Visual inspection of the 95% CIs around each reliability estimate for the preadolescents and early adolescents on the scale and subscales of the NPQC-R indicated that there were overlaps on the Superiority and Exploitativeness subscales, and on the NPQC-R Total scale. In addition, using the Feldt technique, results indicated that there was no bias on the Exploitativeness subscale, F (563, 136) = 1.14, p = .18, d = .10, and on the NPQC-R Total scale, F (563, 136) = 1.12, p = .21, d = .10, between preadolescents and early adolescents. In addition, significant results were found between the two groups on the Superiority subscale, F (563, 136) = 1.38, p = .01, d = .11; however, a negligible effect size suggest that construct bias is not present between the two

groups on the Superiority subscale. Taken together, results do not support construct bias between preadolescents and early adolescents on the Superiority and Exploitativeness subscales, and the NPQC-R Total scale.

Research Question 3

The third research question seeks to examine whether there are differences in narcissism across culture, gender, and age. To address this research question, at least partial strong invariance has to be established between groups. Indeed, findings from research question 1 in this study indicated that strong invariance in the NPQC-R scores was tenable across age (preadolescents, early adolescents), and partial strong invariance in the NPQC-R scores was supported across culture (Singapore, U.S.) and gender (males, females). Researchers have agreed that the comparison of latent means can proceed if at least partial strong invariance between groups is supported (Brown, 2015; Milfont & Fischer, 2010). Because at least partial strong invariance in the NPQC-R scores was found between groups, comparisons of latent means between groups were undertaken. The critical ratio (C.R.) value or the standardized mean difference value, which represents the latent mean differences between the two groups, was obtained (Byrne, 2001).

Research Question 3a. Research question 3a seeks to examine whether there were differences in narcissism between Singapore and U.S. students. Comparisons of latent means across culture were performed by fixing the latent means of the reference group (i.e., Singapore) to zero, and freely estimating the latent means in the comparison group (i.e., U.S.). Results indicated that there were no significant differences between the U.S. and Singapore cultures on the Exploitativeness factor (C.R. = +.17, p = .13). However, on the Superiority factor, significant

results were found between cultures, with the U.S. having a higher latent mean value compared to Singapore (C.R. = +.99, $p < .01$).

Research Question 3b. Research question 3b seeks to examine whether there were differences in narcissism between males and females. Comparisons of latent means across gender were performed with males as the reference group and females as the comparison group. Results indicated that there were no significant differences between males and females on the Superiority factor (C.R. = -.08, $p = .31$), but males were found to have a higher latent mean value on the Exploitativeness factor compared to females (C.R. = -.18, $p = .05$).

Research Question 3c. Research question 3c seeks to examine whether there were differences in narcissism between preadolescents and early adolescents. Comparisons of latent means across age were performed with the preadolescents group serving as the reference group and the early adolescents group serving as the comparison group. Results indicated that there were no significant differences between preadolescents and early adolescents on the Exploitativeness factor (C.R. = +.15, $p = .26$). However, significant differences were found between preadolescents and early adolescents on the Superiority factor (C.R. = -.42, $p < .01$), and this suggests that preadolescents had a higher latent mean value than early adolescents on this factor.

Summary

In conclusion, construct bias in the NPQC-R scores across groups was examined using factor analytic methods. Through this method, partial measurement invariance across culture (Singapore, U.S.) and gender (males, females), and strong measurement invariance between age (preadolescents, early adolescents) were found. In addition, a pilot test conducted to examine measurement invariance across males and females in Singapore and the U.S., found partial

strong invariance across these four subgroups (Singapore males, Singapore females, U.S. males, and U.S. females). Construct bias of the NPQC-R scores across groups using reliability methods also found overlaps in the 95% confidence interval around each internal consistency reliability estimate across groups (i.e., culture – Singapore, U.S., gender – males, females, and age – preadolescents, early adolescents).. Another internal consistency reliability method to examine construct bias was used, the Feldt technique. Results indicated that five out of nine comparisons did not indicate the presence of construct bias across groups. These group comparisons that did not indicate the presence of construct bias included comparisons on the NPQC-R Exploitativeness subscale: Singapore and U.S., males and females, and preadolescents and early adolescents, and the NPQC-R Total scale: Singapore and U.S., and preadolescents and early adolescents. Although significant results were found for the other four group comparisons, i.e., comparisons on the NPQC-R Superiority subscale: Singapore and U.S., males and females, and preadolescents and early adolescents, and group comparison on the NPQC-R Total scale for males and females, negligible effect sizes were reported for all analyses performed involving the Feldt technique, supporting no construct bias in the NPQC-R scores. Overall, the evidence obtained from the factor analytic and reliability methods indicated that there was no construct bias in the NPQC-R scores across culture (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents). Last, latent means comparisons across culture, gender, and age were performed. Results indicated no differences between Singapore and U.S. students on the Exploitativeness factor, but U.S. students had a higher latent mean value than Singapore students on the Superiority factor. Also, males were found to have a higher latent mean value than females on the Exploitativeness factor, but there were no significant differences in latent mean values between the groups on the Superiority factor. In addition, although no significant

differences were found between preadolescents and early adolescents on the Exploitativeness factor, preadolescents had a higher latent mean value than early adolescents on the Superiority factor.

CHAPTER V

Discussion

The present study examined construct bias in the NPQC-R scores using the factor analysis and reliability methods. The factor analysis method entailed performing tests of measurement invariance on the NPQC-R between groups, while the reliability methods involved examining the overlap in the 95% confidence interval of the internal consistency reliability estimates between groups on the NPQC-R, and by using the Feldt technique to determine whether differences in internal consistency reliability estimates on the NPQC-R between groups existed. In addition, because at least partial strong measurement invariance on the NPQC-R was demonstrated between groups, tests of latent means were employed to determine whether group differences on the latent variables (i.e., Superiority and Exploitativeness) were present on the NPQC-R. The research questions will be discussed in greater detail in this chapter.

Is there construct bias in the NPQC-R scores between groups using the factor analysis method?

The first research question in this study sought to examine construct bias on the NPQC-R using the factor analysis method. Tests of measurement invariance were employed across culture (Singapore, U.S.), gender (male, female), and age (preadolescents, early adolescents). Results indicated that the NPQC-R was invariant across groups, with configural, weak, and partial strong measurement invariance found between Singapore and U.S. students and between males and females, and configural, weak, and strong measurement invariance demonstrated between the preadolescents and early adolescents groups.

Measurement Invariance across Culture. Findings of this present study found configural, weak, and partial strong measurement invariance on the NPQC-R among Singapore and U.S. students. Five thresholds associated with three items were not equivalent for Singapore

and U.S. students. Specifically, thresholds 3 and 4 associated with Item 1 (i.e., I always know what I am doing) and Item 10 (i.e., My body looks good) were not equivalent across groups. In addition, threshold 1 associated with Item 12 (i.e., When I am supposed to be punished, I can usually talk my way out of it) was not equivalent between the two groups. The NPQC-R appears invariant between U.S. and Singapore students after these thresholds were released, and this indicates that there is no evidence to suggest construct bias between these groups. Several reasons may possibly explain why these items operate differently for Singapore and U.S. students, and these reasons are described below.

In regard to Item 1 (i.e., I always know what I am doing), a greater level of the latent response is required before U.S. students will endorse the options of "very much like me" or "completely like me". Education is highly valued by Singapore students as it is considered a core component of Confucianism in Chinese societies, including Singapore (Huang & Gove, 2012). Singapore has a very structured educational system that emphasizes the instrumental role of education in maintaining and promoting the country's economic development (Kam & Gopinathan, 1999). In the Singapore education system, students are deliberately placed into streams of different academic pace or tracks when he/she is approximately 10 years old. In 2008, a subject-based banding system was implemented, such that students are recommended to take a specified combination of subjects based on results from school-based examinations when they are in their fourth year of primary education. Students then sit for the high-stakes Primary School Leaving Examination (PSLE) at the end of their sixth year at the primary school level, with results from this examination determining the ability stream or track (i.e., Special, Express, Normal Academic or Normal Technical) that they are placed in (Ministry of Education, Singapore, 2014). For example, students who are placed in the Normal Technical stream for

secondary school are typically provided preparation for vocational and technical training at the Institute of Technical Education (ITE; Teh, 2014). Therefore, in Singapore, students have a clear direction of what they should be doing from an early age. In contrast, the U.S. educational system deemphasizes streaming or tracking of students (Hanushek & Woessmann, 2005). With a less rigid educational structure, there is greater emphasis on exploration, creativity, and personal growth among U.S. students (Kim, 2005). As such, it is plausible that because of the educational system, Singapore students are more likely to endorse the options of "very much like me" or "completely like me" compared to U.S. students when asked about whether they knew what they are doing.

Pertaining to Item 10 (i.e., My body looks good), Singapore students need a higher level of latent response before they endorse the options of "very much like me" or "completely like me". U.S. students are more likely to endorse these options as body image is viewed as an important aspect of their identity and perception of beauty (Frith, Shaw, & Cheng, 2005; Thompson & Hirschman, 1995). In Singapore, however, the body is less emphasized; instead, the face is viewed as being more important. In a study of advertisement magazines in Singapore, Taiwan, and the U.S., female Asian models were depicted in more demure ways than their Caucasian counterparts (Frith et al., 2005). Similarly, Khoo and Karan (2007) found male models in a Singapore magazine to be demurely dressed. As such, it is plausible that Singapore students require a higher level of latent response before they will endorse the options of "very much like me" or "completely like me" on this item.

In addition, Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) operates differently for U.S. and Singapore students, with Singapore students less likely to endorse the option of "not at all like me" compared to U.S. students. This may appear

counterintuitive because Singapore adheres to the tenets of Confucianism, and demonstrating respect for and obedience to one's elders features prominently in the Singapore culture, and would suggest that young individuals would refrain from arguing with adults when they are in the wrong compared to U.S. students (Ho, 1994). However, Ow (2014) asserted that while communication used to be unidirectional from the adult to the child, the younger generation of Singaporeans presently enjoys a more egalitarian relationship with adults due to the influence of Western media and greater educational opportunities. This egalitarian relationship is characterized by mutual respect for each other's opinion (Ow, 2014). In addition, Singapore students tend to avoid being punished, so as to prevent bringing shame to the family. Shame is an emotion that an individual seeks to avoid after being made aware by another individual that he/she has violated a social convention, rule or goal (Fischer & Tangney, 1995). As such, it is possible that Singapore students are more likely to express themselves when they are in trouble due to a more egalitarian relationship with their elders, and their tendency to avoid punishment and bring shame to the family.

After these thresholds for Items 1, 10, and 12 were released, the NPQC-R appeared invariant between Singapore and U.S. students. This finding suggests that there is no evidence of construct bias between these groups.

Measurement Invariance across Gender. When tests of measurement invariance on the NPQC-R across gender were performed, configural, weak, and partial strong measurement invariance on the NPQC-R were found. One threshold associated with Item 2 (i.e., If I ruled the world, it would be a better place) was not equivalent for males and females, and suggest that a greater level of the latent response is required before females will endorse the option of "completely like me". It is plausible that this item may operate differently for males and females,

as males and females recognize that others may perceive their display of dominance differently. Based on the status incongruity hypothesis, females who demonstrate excessive dominance will suffer a backlash from society (Rudman, Moss-Racusin, Glick, & Phelan, 2012). Rudman et al. (2010) mentioned that females who are excessively dominant may have to pay a "dominance penalty", such as by being perceived as having a lack of competence or by being liked less by others. As such, it is possible that females are less likely to endorse the option of "completely like me" on this item as they tend to refrain from demonstrating their dominance, so as to not be perceived in a negative light by others. After this threshold was released, the NPQC-R appeared invariant between males and females, and this suggests that there is no evidence of construct bias between these two groups.

Measurement Invariance across Age. When tests of measurement invariance on the NPQC-R across age were performed, configural, weak, and strong invariance on the NPQC-R were found. These findings suggest that the NPQC-R is invariant across preadolescents and early adolescents.

Other Tests of Measurement Invariance. A pilot test to examine whether the four subgroups (i.e., Singapore males, U.S. males, Singapore females, U.S. females) measured the same construct on the NPQC-R indicated that configural and weak invariance was tenable across all groups. In addition, after freeing the following thresholds specific to the following items, partial strong invariance was tenable across all groups. The thresholds specific to the items that were relaxed included thresholds 3 and 4 on Item 1 (i.e., I always know what I am doing), Item 2 (i.e., If I ruled the world, it would be a better place), and Item 10 (i.e., My body looks good), as well as thresholds 1, 3, and 4 on Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it).

Thresholds 3 and 4 on Item 1 (i.e., I always know what I am doing) were found to be not equivalent across the groups, as a greater level of latent response is required before both U.S. males and females endorse the options of "very much like me" and "completely like me". As mentioned earlier, education is of great importance to Singapore students, and the very structured educational system places students into certain tracks that may determine their future from as young as 10 years of age (Kam & Gopinathan, 1999). In comparison, the U.S. educational system is less rigid and encourages greater exploration of future educational and career paths from elementary through high school (Hanushek & Woessmann, 2005; Kim, 2005). As such, it is plausible that Singapore males and females were more likely to endorse the options of "very much like me" and "completely like me" on Item 1 because their lives are structured around the rigid educational system in their country.

A couple of thresholds on Item 2 (i.e., If I ruled the world, it would be a better place) were not equivalent across groups First, threshold 4 on Item 2 was not equivalent across the groups, as Singapore males were more likely to endorse the option of "completely like me" when compared to all groups. This could be due to the pervasiveness of Confucianism in Singapore's society, in which the patriarchal character of society perpetuates the idea of male dominance and authority (Gupta et al., 2003). In other words, Singapore males are viewed as having greater power and authority in society, so as to provide for and protect his family (Chan, 2000). As such, it is plausible that Singapore males are more likely to endorse the option of "completely like me" on the statement that the world would be a better place if they rule it. In addition, threshold 3 on Item 2 was not equivalent across the groups, as Singapore females were less likely to endorse the option of "very much like me". This may be due to the backlash that woman receive from society if they exhibit excessive dominance, and are perceived as being less liked by others as a result

(Rudman et al., 2010). This may be particularly relevant among Singapore females due to the permeation of Confucian ideology in Singapore's culture, as females are expected to maintain the more traditional gender role of caring for the family and they are valued for this traditional role as compared to being in the role of power and authority (Aryee, 1992; Chan, 2000). Indeed, Wirth (2001) found that Singapore women held only 27% of managerial positions compared to women in the United States where women hold approximately 41.5% of managerial jobs.

On Item 10 (i.e., My body looks good), thresholds 3 and 4 were not equivalent across the groups, as U.S. males were more likely to endorse the options of "very much like me" and "completely like me" on this item as compared to the other groups. During the teen years, males are more likely to be satisfied with their body because pubertal physical changes, such as increase in muscle mass and height, tend to bring males closer to the ideal male masculine body, whereas pubertal changes result in greater body dissatisfaction among females (Jung, Forbes, & Lee, 2009). Yang, Gray and Pope (2005) further explicated that western men tend to focus on their bodies to build their self-esteem regarding their masculinity, whereas Asian men do not base their masculine self-esteem on their body. Indeed, Jung et al. (2009) found that U.S. boys were most satisfied with their appearance when compared to U.S. girls, Korean boys, and Korean girls. As such, it is highly possible that U.S. males would endorse the options of "very much like me" and "completely like me" compared to other groups, so as to boost their masculine self-esteem.

Several thresholds on Item 12 (i.e., When I am supposed to be punished, I can talk my way out of it) were not equivalent across the groups. Thresholds 1, 3, and 4 on Item 12 were not equivalent across the groups, and these indicate that Singapore females were less likely to endorse the options of "not at all like me" compared to other groups, while Singapore males

were more likely to endorse the options of "very much like me" and "completely like me" compared to other groups. This alludes to Singapore students, especially Singapore males, being more likely to talk their way out of punishment compared to U.S. students. As mentioned previously, Singapore students, both males and females, tend to avoid being punished, so as to prevent bringing shame to the family. This is particularly evident among Singapore males as they seek to avoid being shamed, in order to protect their masculinity in Singapore's patriarchal society (Tamney, 1995). In fact, males, especially Singapore males, tend to have more self-stigma as they need to meet societal expectations of being an authority figure or in a position of authority (Gupta et al., 2003; Vogel, Wade, & Haake, 2006). As such, it is highly possible that Singapore students, especially Singapore males, are more likely talk their way out of punishment when they are in trouble in order to avoid punishment and shame when compared to the other groups.

Is there construct bias in the NPQC-R scores between groups using the reliability methods?

The second research question sought to examine construct bias on the NPQC-R using the reliability methods by comparing the overlap in the 95% confidence intervals of the internal consistency reliability estimates across groups, as well as by using the Feldt technique. Results indicated that there was overlap in the 95% confidence intervals of the internal consistency reliability estimates on the scale and subscales of the NPQC-R between groups. In addition, five out of nine comparisons were not significant when the Feldt technique was used to test whether differences in the internal consistency reliability estimates were present between groups. However, when effect sizes were computed for group differences, negligible effect sizes were found between groups on the scale and subscales of the NPQC-R; thus, this would suggest that construct bias is not present between groups on the scale and subscales of the NPQC-R. Taken

together, a preponderance of evidence using both reliability methods suggests that overall the NPQC-R does not demonstrate construct bias between groups. Specifically, no construct bias was found between Singapore and U.S. students, males and females, and preadolescents and early adolescents using the internal consistency reliability approaches.

Are there differences in narcissism between groups?

The third research question sought to examine whether there were differences between groups on the NPQC-R. Specifically, this research question seeks to identify if there are cultural (Singapore, U.S.), gender (males, females), and age (preadolescents, early adolescents) differences on the NPQC-R. As at least partial strong measurement invariance was found across culture, gender, and age, means on the latent variables between groups were compared. In other words, differences in groups were examined to determine whether true differences on the latent variables were present. By using the test of latent means approach to examine differences in groups, results are viewed with greater confidence than the observed mean score difference approach, as the latent means approach removes measurement error related to the latent variable (Dimitrov, 2006).

Cultural Differences. Although Fung and colleagues (2010) found higher levels of overall narcissism among Hong Kong adolescents compared to U.S. adolescents, this is the first study to explore whether differences across dimensions of narcissism between students of different cultural groups exist. Findings from this study indicated that U.S. students had higher levels of superiority than Singapore students. Based on Hofstede's (1991) classification of the U.S. being considered an individualistic culture and Singapore being more collectivistic, this result is relatively similar to results from a large scale study conducted by Foster et al. (2003) on narcissism levels between individualistic and collectivistic cultures across the world. In Foster

and colleagues' (2003) study, individualistic cultures, such as the U.S. and Canada, were found to exhibit higher levels of superiority when compared to collectivistic cultures, such as cultures in the Asia and the Middle East regions. One reason for the elevated levels of superiority among U.S. students compared to Singapore students may be that features of superiority may be more aligned with characteristics of boastful self-presentation, and research has found that this self-promotion style is more prevalent and acceptable in individualistic cultures compared to collectivistic cultures (Chen & Jing, 2012).

In regard to exploitativeness, similar levels of exploitativeness were found between the Singapore and U.S. students. Similar levels of exploitativeness between these two cultures may be inferred from studies examining the relations between exploitaitveness and proactive aggression, as these studies have consistently demonstrated positive relations between exploitativeness and proactive aggression (Ang & Raine, 2009; C. Barry, Grafeman et al., 2007; T. Barry et al., 2007; Salmivalli, 2001; Washburn et al., 2004). A recent cross-cultural study by Ang, Huan, and Florell (2014) found similar levels of proactive aggression between U.S. and Singapore adolescents. Thus, it may be inferred that similar levels of exploitativeness are exhibited by U.S. and Singapore students. In addition, although this finding may appear counterintuitive as Singapore, being a collectivistic culture, should have lower levels of exploitativeness as compared to the U.S., which is an individualistic culture, similar levels of exploitativeness in both cultures could possibly be explained by similar levels of economic development in both cultures. Between 2010 and 2014, Singapore's average gross domestic product (GDP) per capita of $55,182.50 was relatively similar to the GDP per capita of $53,042.00 in the U.S. (The World Bank Group, 2015). Because both the U.S. and Singapore are wealthy countries, Triandis (1989) asserted that it is plausible that individuals from the U.S. and

Singapore tend to pursue their self-interest in an effort to accumulate more wealth. In particular, in the Singapore context, the initiation of the Global City project has resulted in an influx of foreigners into this small nation with limited resources (Oswin & Yeoh, 2010). Owing to this, Singapore students may develop a competitive streak and seek to gain an advantage over others so as to survive in the Singapore society. As such, the aforementioned reasons may explain similar levels of exploitativeness between Singapore and U.S. students.

Gender Differences. This study found similar levels of superiority between males and females, and concurs with findings from Fossati et al.'s (2010) study on adolescent males and females. Studies (e.g., Twenge & Foster, 2008; Twenge et al., 2008) have found that narcissistic characteristics, including characteristics of superiority, have increased over time, with more elevated scores among students in 2005-2006 compared to 1985-1989. Twenge (2008) further asserted that this increase is prevalent across both males and females. It is possible that similar levels of superiority could be due to similar parenting practices among males and females nowadays, with parents lavishing excessive praise on their children. Research has found that it is common for parents to deliver inflated praise to children, and this did not differ between males and females (Brummelman, Thomaes, de Castro, Overbeek, & Bushman, 2014). Excessive praise lavished on children may result in a false self-image, such that children develop distorted cognitions about the self and regard themselves as being more superior to others (Baumeister & Vohs, 2001; Millon, 1981). A recent study agreed that the development of narcissism could be attributed to children internalizing inflated parental beliefs that they are special or entitled (Brummelman et al., 2015).

In the area of exploitativeness, males were found to exhibit higher levels of exploitativeness than females. This result is in contrast to previous studies conducted on

adolescents (Fossati et al., 2010; Washburn et al., 2004). It is possible that this discrepant finding could be due to the different samples used in their studies. Fossati et al's (2010) study consisted of Italian high school students, whereas Washburn et al.'s (2004) study included predominantly African American students. However, the finding that males displayed higher levels of exploitativeness than females in the current study largely mirrors results obtained from a large scale study conducted on college students, in which males demonstrated more exploitativeness/entitlement characteristics than females (Tschanz et al., 1998). Moreover, Grijalva et al.'s (2014) recent meta-analysis results concurred that males displayed more exploitativeness/entitlement characteristics than females. Tschanz and colleagues (1998) suggested that females have lower levels of exploitativeness and entitlement than males because characteristics of exploitativenes and entitlement are less well-integrated in females than males. In other words, exploitative behaviors go against expectations that guide how females should behave. Indeed, Bradlee and Emmons (1992) found that exploitativeness/entitlement was negatively related to communal characteristics that are commonly associated with females, such as nurturance and friendliness.

Age Differences. Findings from this study found higher levels of superiority among preadolescents than early adolescents, and no differences in exploitativeness were found between these two groups. As narcissism has been found to be related to the development of adolescent egocentrism (Aalsma et al., 2006; Alberts et al., 2007), it is puzzling that results from this study were in contrast to this theory. It is plausible that this may be due to how students were categorized for the current study. As children and adolescents develop adolescent egocentrism differently, it may be more useful to group students based on their development of adolescent egocentrism.

Limitations

There are several limitations associated with this study. These include the small sample size in some subgroups, the exploratory nature of partial measurement invariance, the use of a convenience sample, the limited age range in the sample, the use of a self-report measure, the time lag in data collection in the two cultures, and the reliance on only three methods to examine construct bias in the NPQC-R scores.

First, the small sample sizes for some subgroups in this study could limit the statistical power to detect model misspecification, and to detect differences in groups in the present study (Brown, 2015). In addition, a small sample size may limit the precision in the measurement invariance model, such that the sample data and model may not approximate the population. In this present study, the sample size of U.S. males and U.S. females were 73 and 78 participants, respectively. Muthén recommended that analyses with medium-sized models of between 10 and 15 indicators should be performed using sample sizes of between 150 and 200 participants (Brown, 2015). As such, to ensure that the model has adequate power and precision, replication of the present study with a larger sample size for each subgroup, especially for U.S. males and U.S. females, is needed.

Second, while measurement invariance analyses performed between groups resulted in at least partial strong measurement invariance between groups, Brown (2015) stated that due to parameters being freely estimated between groups, partial measurement invariance is exploratory in nature. As such, cross-validation using an independent sample should be performed in order to confirm the presence of measurement invariance on the NPQC-R between groups (Byrne et al, 1989).

Third, archival data analyzed in this study was collected using a convenience sample. A convenience sample limits the ability to generalize findings from this study to the general population. Indeed, the demographic characteristics, such as the ethnic composition, of both the Singapore and U.S. subsamples are not representative of the general population. For example, 60.5% of participants in the Singapore sample self-identified as being Chinese, whereas 74.3% of the Singapore population are of Chinese ethnicity (Singapore Department of Statistics, 2014); in addition, while Hispanics make up 17.4% of the U.S. population (Colby & Ortman, 2015), only 2.0% in the U.S. sample self-identified as being Hispanic. To ensure that findings are generalizable to the general population, future studies should obtain a sample that is representative of the general population. One suggested method would be to use a proportionate stratified sampling method (Engel & Schutt, 2013). In this method, sampling strata based on a particular demographic characteristic, such as ethnicity, are created. Participants are then randomly selected from each stratum in the exact proportion that is representative of the general population.

Fourth, participants in this study were limited to children and adolescents between the ages of 12 and 14. Due to this limited age range, inferences made from findings of this study are restricted to children and adolescents between 12 and 14 years old, and extension of findings from this study to children and adolescents of other ages cannot be made. To better understand construct bias in the NPQC-R during the developmental process, and the similarities or differences in narcissistic personality characteristics across the developmental lifespan, future studies may include data collected across a wider age range, such as from middle childhood to late adolescence. Furthermore, for a more nuanced analysis into narcissism between

preadolescents and early adolescents, it is recommended that students are grouped according to their adolescent egocentrism development.

Fifth, as the NPQC-R is a self-report measure, there may be concerns about participants' truthfulness in responses. It is plausible that some participants may respond in a socially desirable manner, so as to present themselves in a favorable light (McIntire & Miller, 2007). Furthermore, in some cultures, especially East Asian cultures that are influenced by Confucian values, participants may be more modest in self-presentation to avoid standing out from the group, such that their responses may not be reflective of their innermost feelings (Kitayama, Markus, Matsumoto, & Norasakkunit, 1997). As there may be inaccuracies in self-reports, Bolkan and Hooker (2012) suggested that obtaining both self-reports and ratings from others may be the best approach in assessing an individual's personality characteristic. Indeed, Hogan (1998) asserted that since personality is a multi-faceted concept that entails self- and observer-perspectives, obtaining ratings from other informants, such as parents, teachers, and peers, in addition to self-reports may provide a more accurate description of an individual's personality characteristics.

Sixth, there is a time lag in data collection between the Singapore and U.S. student subsamples, with data from the Singapore student subsample collected several years prior to data collected from the U.S. student subsample. Due to this time lag in data collection, it is plausible that cohort effects, which is defined as characteristics among individuals who share similar temporal experiences (Di Maggio, Martinez, Menard, Petit, & Thibuaut, 2001), may impact the results from this study. Although the ages of participants in both samples are similar, the shared experiences during the time period that data were collected in the Singapore and U.S. subsamples may be different. For example, when data were collected in the Singapore subsample, social

networking sites, such as Facebook, just started to emerge, and only featured basic functions; however, when data were collected in the U.S., there was a proliferation of social networking sites, such as Twitter and Instagram, which allow users to share information about themselves with others using numerous means. Studies have found that the use of social networking sites has been found to be related to narcissism, such that individuals with narcissism are more likely to promote themselves to others using social networking sites (e.g., Buffardi & Campbell, 2008). As such, to ensure that cohort effects do not confound results in future studies, concurrent collection of data from the two samples is highly recommended.

Last, although this study utilized three methods to examine construct bias, namely measurement invariance across groups, the overlap in the 95% confidence intervals of the internal consistency reliability estimates between groups, and the Feldt technique, there are still other methods to examine construct bias. Indeed, Reynolds and Suzuki (2012) recommended that diverse methods be used to assess for bias, so as to obtain reliable evidence of bias. Another reliability method used to examine for construct bias includes the test-retest reliability method. In this method, test-retest correlation coefficients in the NPQC-R across groups are tested. If no differences are found across groups, this would suggest that no bias in the NPQC-R is present between the groups. As such, future studies should look into corroborating information from this test-retest reliability method with other methods used to examine construct bias, in order to determine whether construct bias in the NPQC-R exists across groups.

Research Implications and Future Research Directions

This dissertation responded to the call by Reynolds and Lowe (2009) for more studies to use objective and psychometrically sound methods to examine test bias in personality tests due to a dearth of studies in this area; as such, this dissertation makes a novel contribution to the extant

literature by examining construct bias in the NPQC-R scores across culture, gender, and age. Specifically, this is the first study to examine construct bias in the NPQC-R scores using both the factor analysis and reliability methods across culture (Singapore, U.S.), gender (males, females), and age (preadolescence, early adolescence). As results from this study did not suggest the presence of construct bias in the NPQC-R scores across groups, this would indicate that the NPQC-R assesses a similar underlying construct across groups. This suggests that it is plausible to use the NPQC-R to develop theories across these respective groups in future research. Along the same vein, this current study portends numerous areas for potential future research. Future research directions may include exploring other sources of bias on the NPQC-R, exploring potential variables that may explain group differences, and exploring invariance of the NPQC-R with other subgroups.

One area of future research would be to explore other sources of bias. Reynolds and Suzuki (2012) asserted that there are three main types of test bias, i.e., construct bias, predictive bias, and content bias, and that it is pertinent that all three sources of test bias are examined before a test is published. As this study only examined construct bias in the NPQC-R, future research studies should examine predictive and content bias on the NPQC-R. Predictive bias refers to bias in inferences of future outcomes or behaviors that can be made from test scores (Reynolds & Lowe, 2009). This entails obtaining a regression line between narcissistic characteristics and an outcome variable for each group. If no predictive bias is found between groups, the groups should have similar intercepts and/or regression slopes. A common statistical method used to test for predictive bias is Potthoff's (1966) method. In this method, regression coefficients and intercepts across groups are tested concurrently. Should a non-significant result be found, this would indicate that there is no predictive bias in the NPQC-R between groups;

conversely, if a significant result is found, separate tests for slopes and intercepts may be conducted to determine how well the NPQC-R overestimates or underestimates the prediction of an outcome variable for each group. Content bias refers to whether items on the test measure all aspects of the construct. Reynolds and Suzuki (2012) mentioned that cultural factors, such as cultural beliefs about aggression, may influence personality tests. To detect whether content bias is present in a measure, tests at the item level are performed. One technique described by Reynolds, Lowe, and Saenz (1999) is the partial correlation method. Reynolds et al. (1999) stated that this method is a powerful method that allows for detecting if item bias is present. In this method, a partial correlation between an item score and a demographic variable is computed, after partialing out the correlation between the demographic variable and the total test score. If the demographic variable and item score is still correlated after removing the partial correlation, this indicates that there is item bias, and that the item performs differently between the groups (Reynolds et al., 1999; Reynolds & Lowe, 2009). Taken together, by examining all sources of bias among groups and finding no bias among groups, this will ensure that the interpretation of NPQC-R scores across groups is valid.

Another area that could be examined in future research is to explore potential variables that may explain the differences in narcissism between groups. As one of the aims of this study was to determine whether differences between groups on the NPQC-R were present, data on explanatory variables to explicate why group differences exist were not collected as it was beyond the scope of this study. Future research could look into variables such as degree of collectivism, degree of adherence to gender norms, and degree of adolescent egocentrism development to explain similarities or differences in narcissism.

Future research may also examine cultural validity of the NPQC-R across Singapore and U.S. students. Cultural validity seeks to examine how socio-cultural factors, e.g., the values and beliefs of individuals, impact how these individuals think, interpret, and respond to items on a measure (Solano-Flores & Nelson-Barber, 2001). Although the NPQC-R was found to have similar *meaning* between Singapore and U.S. students as weak measurement invariance was tenable between these groups, Solano-Flores (2011) asserted that cultural validity is a critical concept that still needs to be examined when measures are used across cultures. As such, to determine whether socio-cultural experiences influence how students in both groups make sense of items on the NPQC-R, examining cultural validity of the NPQC-R across these groups is highly recommended.

Last, future studies should explore invariance of the NPQC-R scores among other subgroups. Since narcissism has been found to be closely related to mental health conditions, such as conduct problems, it would be interesting to examine whether the NPQC-R scores are invariant across children and adolescents with and without mental health disorders. In addition, personality traits have rarely been examined among children and adolescents with disabilities (e.g., Sharma, 2004). In particular, there are currently no studies that examined narcissistic personality characteristics in students with and without disabilities. As such, future studies should look into whether the NPQC-R scores are invariant among children and adolescents with and without disabilities.

Practical Implications

A major practical implication of this study is that practitioners, including school psychologists, will be more aware that the NPQC-R scale and subscales function similarly across culture (Singapore, U.S.), gender (male, female), and age (preadolescents, early adolescents). In

other words, the narcissism construct on the NPQC-R are similar between groups. By being aware of the psychometric properties of the NPQC-R and the notion that there was no indication of the presence of construct bias between groups, practitioners should be more confident in using the NPQC-R to measure narcissism in the aforementioned groups. Specifically, practitioners can be assured that an individual's score on the NPQC-R reflects his/her "true score" and is not due to his/her group membership. In addition, because no construct bias was found between groups on the NPQC-R, comparisons between these groups using the NPQC-R to measure narcissistic characteristics can be made across these groups.

Conclusion

Overall, this study demonstrated that there was no evidence of construct bias in the NPQC-R scores across culture (Singapore, U.S.), gender (male, female), and age (preadolescents, early adolescents). This conclusion was drawn based on the preponderance of evidence obtained from results found through the factor analytic method, i.e., at least partial strong measurement invariance between groups, and through reliability methods, i.e., overlap in the 95% confidence intervals of the internal consistency reliability estimates on the NPQC-R between groups, and the Feldt technique to determine whether differences in internal consistency reliability estimates were present in the groups. Furthermore, tests of latent means between cultural groups indicated elevated levels of Superiority in U.S. students compared to Singapore students, but there were no differences in the levels of Exploitativeness between the two groups. When latent means were tested between gender, males were found to exhibit higher levels of Exploitativeness than females, but similar levels of Superiority were found between. In addition, the preadolescent group demonstrated higher levels of Superiority compared to the early adolescents group, but no differences were found between these two groups in the levels of

Exploitativeness. Notwithstanding the limitations of this study, this study supports the use of the NPQC-R in future research and in clinical and school settings among Singapore and U.S. students, males and females, and preadolescents and early adolescents.

References

Aalsma, M. C., Lapsley, D. K. & Flannery, D. J. (2006). Narcissism, personal fables, and adolescent adjustment. *Psychology in the Schools. 43,* 481-491. doi:10.1002/pits.20162

Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R.W., & Kashy, D. A. (2011). What does the Narcissistic Personality Inventory really measure?, *Assessment, 18,* 67–87. doi:10.1177/1073191110382845

Alberts, A., Elkind, D., & Ginsberg, S. (2007). The personal fable and risk-taking in early adolescence. *Journal of Youth and Adolescence, 36,* 71-76. doi:10.1007/s10964-006-9144-4

Allison, P. D. (2009). Missing data. In R. E. Millsap & A. Maydeu-Olivares (Eds.*), The SAGE handbook of quantitative methods in psychology* (pp. 73-91). London, United Kingdom: SAGE. doi:10.4135/9780857020994.n4

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of Research of Personality, 40,* 440–450. doi:10.1016/j.jrp.2005.03.002

Ang, R. P., Huan, V. S., & Florell, D. (2014). Relationship between proactive and reactive aggression, and cyberbullying across United States and Singapore adolescent samples. *Journal of Interpersonal Violence, 29,* 237-254. doi:10.1177/0886260513505149

Ang, R. P., Ong, E. Y. L., Lim, J. C. Y., & Lim, E. W. (2010). From narcissistic exploitativeness to bullying behavior: The mediating role of approval-of-aggression beliefs. *Social Development, 4,* 721-735. doi:10.1111/j.1467-9507.2009.00557.x

Ang, R. P., & Raine, A. (2009). Reliability, validity, and invariance of the Narcissistic Personality Questionnaire for Children – Revised (NPQC-R). *Journal of Psychopathology and Behavior Assessment, 31,* 143-151. doi:10.1007/s10862-008-9112-2

Ang, R. P., Tan, K. A., & Mansor, A. T. (2011). Normative beliefs about aggression as a mediator of narcissistic exploitativeness and cyberbullying. *Journal of Interpersonal Violence, 26,* 2619–2634. doi:10.1177/0886260510388286

Ang, R. P., & Yusof, N. (2006). Development and initial validation of the Narcissistic Personality Questionnaire for Children: A preliminary investigation using school-based Asian samples. *Educational Psychology, 26,* 1–18. doi:10.1080/01443410500340942

Ansell, E. B. (2005). *The structure of narcissistic personality: Adaptive and maladaptive dimensions as an integrated model of narcissism* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Theses database. (UMI No. 3292490)

Aryee, S. (1992). Antecedents and outcomes of work-family conflict among married professional women: Evidence from Singapore. *Human Relations, 45,* 813-837. doi:10.1177/001872679204500804

Ashby, H. U., Lee, R. R., & Duke, E. H. (1979, September). *A narcissistic personality disorder MMPI scale.* Paper presented at the 87th Annual Convention of the American Psychological Association, New York, NY.

Bachar, E., Hadar, H., & Shalev, A. Y. (2005). Narcissistic vulnerability and the development of PTSD: A prospective study. *Journal of Nervous and Mental Disease, 193,* 762-765. doi:10.1097/01.nmd.0000185874.31672.a5

Banai, E., Mikulincer, M., & Shaver, P. R. (2005). "Selfobject" needs in Kohut's self psychology: Links with attachment, self-cohesion, affect regulation, and adjustment. *Psychoanalytic Psychology, 29,* 224-260. doi:10.1037/0736-9735.22.2.224

Bandura, A. (2002). Social cognitive theory in cultural context. *Applied Psychology: An International Review, 51,* 269-290. doi:10.1111/1464-0597.00092

Barry, C. T., & Ansel, L. L. (2011). Assessment of youth narcissism. In W. K. Campbell & J. D. Miller (Eds.), *The handbook of narcissism and Narcissistic Personality Disorder: Theoretical approaches, empirical findings, and treatments* (pp. 153-163). Hoboken, NJ: John Wiley & Sons.

Barry, C. T., Frick, P. J., Adler, K. A., & Grafeman, S. J. (2007). The predictive utility of narcissism among children and adolescents: Evidence for a distinction between adaptive and maladaptive narcissism. *Journal of Child and Family Studies, 16,* 508–521. doi:10.1007/s10826-006-9102-5

Barry, C. T., Frick, P. J., & Killian, A. L. (2003). The relation of narcissism and self-esteem to conduct problems in children: A preliminary investigation. *Journal of Clinical Child and Adolescent Psychology, 32,* 139–152. doi:10.1207/s15374424jccp3201_13

Barry, C. T., Grafeman, S. J., Adler, K. A., & Pickard, J. D. (2007). The relations among narcissism, self-esteem, and delinquency in a sample of at-risk adolescents. *Journal of Adolescence, 30,* 933– 942. doi:10.1016/j.adolescence.2006.12.003

Barry, C. T., Pickard, J. D., & Ansel, L. L. (2009). The associations of adolescent invulnerability and narcissism with problem behaviors in at-risk adolescents. *Personality and Individual Differences, 47,* 577-582. doi:10.1016/j.paid.2009.05.022

Barry, C. T., & Wallace, M. T. (2010). Current considerations in the assessment of youth narcissism: Indicators of pathological and normative development. *Journal of Psychopathology and Behavioral Assessment, 32,* 479-489. doi:10.1007/s10862-010-9188-3

Barry, T. D., Thompson, A. H., Barry, C. T., Lochman, J. E., Adler, K., & Hill, K. (2007). The importance of narcissism in predicting proactive and reactive aggression in moderately to highly aggressive children. *Aggressive Behavior, 33,* 185-197. doi:10.1002/ab.20198

Baumeister, R. F., & Vohs, K. D. (2001). Narcissism as addiction to esteem. *Psychological Inquiry, 12,* 206-209.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246. doi:10.1037//0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-600. doi 10.1037//0033-2909.88.3.588

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: John Wiley.

Bolkan, C., & Hooker, K. (2012). Self-regulation and social cognition in adulthood: The gyroscope of personality. In M. Sliwinski and S. Whitbourne (Eds.) *Handbook of developmental psychology: Adult development and aging* (pp. 357-380). New York, NY: Wiley-Blackwell Publishers.

Bond, M. H., Kwan, V. S. Y., & Li, C. (2000). Decomposing a sense of superiority: The differential social impact of self-regard and regard-for-others. *Journal of Research in Personality, 34,* 537-553. doi:10.1006/jrpe.2000.2297

Bontempo, D. E., & Hofer, S. M. (2007). Assessing factorial invariance in cross-sectional and longitudinal studies. In. A. D. Ong & M. H. H. van Dulmen (Eds), *Oxford handbook of methods in positive psychology* (pp. 153-175). New York, NY: Oxford University Press.

Borsboom, D. (2006). When does measurement invariance matter? *Medical Care, 44,* S176-S181. doi:10.1097/01.mlr.0000245143.08679.cc

Bradlee, P. M., & Emmons, R. A. (1992). Locating narcissism within the interpersonal circumplex and the five-factor model. *Personality and Individual Differences, 13,* 821-830. doi:10.1016/0191-8869(92)90056-u

Brown, R. P., Budzek, K., & Tamborski, M. (2009). On the meaning and measure of narcissism. *Personality and Social Psychology Bulletin, 35,* 951-964. doi:10.1177/0146167209335461

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* (2nd ed.). New York, NY: Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Brummelman, E., Thomaes, S., Orobio de Castro, B., Overbeek, G., & Bushman, B. J. (2014). "That's not just beautiful—that's incredibly beautiful!": The adverse impact of inflated praise on children with low self-esteem. *Psychological Science, 25,* 728–735. doi:10.1177/0956797613514251

Buffardi, L. E., & Campbell, W. K. (2008). Narcissism and social networking web sites. *Personality and Social Psychology Bulletin, 34,* 1303-1314. doi:10.1177/0146167208320061

Buss, D. M., & Chiodo, L. M. (1991). Narcissistic acts in everyday life. *Journal of Personality, 59,* 179–215. doi:10.1111/j.1467-6494.1991.tb00773.x

Byrne, B. M. (1994). Testing for factorial validity, replication, and invariance of a measuring instrument: A paradigmatic application on the Maslach Burnout Inventory. *Multivariate Behavioral Research, 26,* 583-605. doi:10.1207/s15327906mbr2903_5

Byrne, B. M., Shavelson, R. J., Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105,* 456-466. doi:10.1037/0033-2909.105.3.456

Cai, H., & Gries, H. P. (2013). National narcissism: Internal dimensions and international correlates. *Psych Journal, 2,* 122-132. doi:10.1002/pchj.26

Cain, N. M., Pincus, A. L., & Ansell, E. B. (2008). Narcissism at the crossroads: Phenotypic description of pathological narcissism across clinical theory, social/personality psychology, and psychiatric diagnosis. *Clinical Psychology Review, 28,* 638-656. doi:10.1016/j.cpr.2007.09.006

Calhoun, G. B., Glaser, B. A., Stefurak, T., & Bradshaw, C. P. (2000). Preliminary validation of the Narcissistic Personality Inventory - Juvenile Offender. *International Journal of Offender Therapy and Comparative Criminology, 44,* 564–580. doi:10.1177/0306624x00445004

Campbell, W. K., Brunell, A. B., & Finkel, E. J. (2006). Narcissism, interpersonal self-regulation, and romantic relationships: An agency model approach. In K. D. Vohs, & E. J. Finkel (Eds.), *Self and relationships: Connecting intrapersonal and interpersonal processes* (pp. 57-83). New York, NY: Guilford.

Campbell, W. K., & Foster, J. D. (2007). The narcissistic self: Background, an extended agency model, and ongoing controversies. In C. Sedikides & S. Spencer (Eds.), *Frontiers in social psychology: The self* (pp. 115–138). Philadelphia, PA: Psychology Press.

Carlson, K. S., & Gjerde, P. F. (2009). Preschool personality antecedents of narcissism in adolescence and emergent adulthood: A 20-year longitudinal study. *Journal of Research in Personality, 43,* 570-578. doi:10.1016/j.jrp.2009.03.003

Chan, J. S. (2000). The status of women in a patriarchal state: The case of Singapore. In L. Edwards & M. Roces (Eds.), *Women in Asia: Tradition, modernity and globalisation* (pp. 39-58). Ann Arbor, MI: The University of Michigan Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464-504. doi:10.1080/10705510701301834

Chen, F. F., & Jing, Y. (2012). The impact of individualistic and collectivistic orientation on the judgment of self-presentation. *European Journal of Social Psychology, 42,* 470-481. doi:10.1002/ejsp.1872

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233-255. doi:10.1207/s15328007sem0902_5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the U.S. population: 2014 to 2060.* Retrieved from U.S. Census Bureau website: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf

Cooper, A. M. (2005). *The quiet revolution in American psychoanalysis: Selected works of Arnold M. Cooper*. New York, NY: Brunner-Routledge.

Corry, N., Merritt, R. D., Mrug, S., & Pamp, B. (2008). The factor structure of the Narcissistic Personality Inventory. *Journal of Personality Assessment, 90,* 593–600. doi:10.1080/00223890802388590

Crockatt, P. (2006). Freud's 'On narcissism: an introduction'. *Journal of Child Psychotherapy, 32,* 4-20. doi:10.1080/00754170600563638

Dadds, M. R., Fraser, J., Frost, A., & Hawes, D. (2005) Disentangling the underlying dimensions of psychopathy and conduct problems in childhood: A community study. *Journal of Consulting and Clinical Psychology, 73,* 400-410. doi:10.1037/0022-006x.73.3.400

Di Maggio, C., Martinez, M., Menard, J. F., Petit, M., & Thibaut, F. (2001). Evidence of a cohort effect for age at onset of schizophrenia. *American Journal of Psychiatry, 158,* 489-492. doi:10.1176/appi.ajp.158.3.489

Dimitrov, D. M. (2006). Comparing groups on latent variables: A structural equation modeling approach. *Work: Journal of Prevention, Assessment & Rehabilitation, 26,* 429-463.

Dong, L., Wu, H., & Waldman, I. D. (2014). Measurement and structural invariance of the Antisocial Process Screening Device. *Psychological Assessment, 26,* 598-608. doi:10.1037/a0035139

Dowson, J. H. (1992). DSM-III-R narcissistic personality disorder evaluated by patients' and informants' self-report questionnaires: Relationships with other personality disorders and a sense of entitlement as an indicator of narcissism. *Comprehensive Psychiatry, 33,* 397–406. doi:10.1016/0010-440x(92)90062-u

Elkind, D. (1967). Egocentrism in adolescence. *Child Development, 38,* 1025–1034. doi:10.1111/j.1467-8624.1967.tb04378.x

Emmons, R. A. (1984). Factor analysis and construct validity of the Narcissistic Personality Inventory. *Journal of Personality Assessment, 48,* 291-300. doi:10.1207/s15327752jpa4803_11

Emmons, R. A. (1987). Narcissism: Theory and measurement. *Journal of Personality and Social Psychology, 52,* 11-17. doi:10.1037//0022-3514.52.1.11

Emond, A., Ormel, J., Veenstra, R., & Oldehinkel, A. J. (2007). Preschool behavioral and social-cognitive problems as predictors of (pre)adolescent disruptive behavior. *Child Psychiatry and Human Development, 38,* 221-236. doi:10.1007/s10578-007-0058-5

Engel, R. J., & Schutt, R. K. (2013). *The practice of research in social work* (3rd ed.). Thousand Oaks, CA: Sage.

Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika, 34,* 363-373. doi:10.1007/bf02289364

Fischer, K. W., & Tangney, J. P. (1995). Self-conscious emotions and the affect revolution: Framework and overview. In J. P. Tangney & K. W. Fischer (Eds.), *Self-conscious emotions: The psychology of shame, guilt, embarrassment, and pride* (pp.3-24). New York, NY: Guilford.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9,* 466-491. doi:10.1037/1082-989x.9.4.466

Fossati, A., Borroni, S., Eisenberg, N., & Maffei, C. (2010). Relations of proactive and reactive dimensions of aggression to overt and covert narcissism in nonclinical adolescents. *Aggressive Behavior, 35,* 1-7. doi:10.1002/ab.20332

Foster, J. D., & Campbell, W. K. (2007). Are there such things as "narcissists" in social psychology? A taxometric analysis of the Narcissistic Personality Inventory. *Personality and Individual Differences, 43,* 1321-1332. doi:10.1016/j.paid.2007.04.003

Foster, J. D., Campbell, W. K., & Twenge, J. M. (2003). Individual differences in narcissism: Inflated self-views across the lifespan and around the world. *Journal of Research in Personality, 37,* 469-486. doi:10.1016/s0092-6566(03)00026-6

Frick, P., Bodin, S. D., & Barry, C. T. (2000). Psychopathic traits and conduct problems in community and clinic-referred samples of children: Further development of the Psychopathy Screening Device. *Psychological Assessment, 12,* 382-393. doi:10.1037/1040-3590.12.4.382

Frick, P. J., O'Brien, B. S., Wootton, J. M., & McBurnett, K. (1994). Psychopathy and conduct problems in children. *Journal of Abnormal Psychology, 103,* 700-707. doi:10.1037//0021-843x.103.4.700

Frith, K., Shaw, P., & Cheng, H. (2005). The construction of beauty: A cross-cultural analysis of women's magazine advertising. *Journal of Communication, 55,* 56-70. doi:10.1111/j.1460-2466.2005.tb02658.x

Fukunishi, I., Nakagawa, T., Nakamura, H., Li, K., Hua, Z. Q., & Krantz, T. S. (1996). Relationships between type A behavior, narcissism, and maternal closeness for college students in Japan, the United States of America, and the People's Republic of China. *Psychological Reports, 78,* 939–944. doi:10.2466/pr0.1996.78.3.939

Fung, A., Gao, Y., & Raine, A. (2010). The utility of the child and adolescent psychopathy construct in Hong Kong, China. *Journal of Clinical Child and Adolescent Psychology, 39*, 134-140. doi:10.1080/15374410903401138

Furr, M. R., & Bacharach, V. R. (2013). *Psychometrics: An introduction* (2nd ed.). Thousand Oaks, CA: SAGE.

Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. K. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory–13 and the Narcissistic Personality Inventory-16. *Psychological Assessment, 25*, 1120-1136. doi:10.1037/a0033192

Ginevra, M. C., Nota, L., Soresi, S., Shogren, K. A., Wehmeyer, M. L., & Little, T. D. (2013). A cross-cultural comparison of the self-determination construct in Italian and American adolescents. *International Journal of Adolescence and Youth,* Advance online publication. doi:10.1080/02673843.2013.808159.

Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015). Gender differences in narcissism: A meta-analytic review. *Psychological Bulletin, 141,* 261-310. doi:10.1037/a0038231

Gupta, M. D., Jiang, Z., Li, B., Xie, Z., Chung, W., & Bae, H. (2003). Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *The Journal of Development Studies, 40,* 153-187. doi:10.1080/00220380412331293807

Hanushek, E. A., & Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *The Economic Journal, 116,* C63-C76. doi:10.1111/j.1468-0297.2006.01076.x

Harter, S. (2006). The self. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology (Vol. 3). Social, emotional, and personality development* (6th ed., pp. 505–570). New York, NY: Wiley.

Hathaway, S. R., & McKinley, J. C. (1951). *The MMPI manual.* New York, NY: Psychological Corporation.

Hendin, H. M., & Cheek, J. M. (1997). Assessing hypersensitive narcissism: A reexamination of Murray's Narcism Scale. *Journal of Research in Personality, 31,* 588-599. doi:10.1006/jrpe.1997.2204

Ho, D. Y. (1994). Filial piety, authoritarian moralism, and cognitive conservatism in Chinese societies. *Genetic, Social, and General Psychology Monographs, 120,* 347-365.

Hoffman, L. (2014, January 1). Multiple-group measurement invariance in IFA using Mplus [Video file]. Retrieved from http://camrelay1.unl.edu/inbox/lhoffman2/948_Lecture9_Part3_Example9_-_20140414_114948_7.html

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. London, United Kingdom: McGraw-Hill.

Hogan, R. (1998). Reinventing personality. *Journal of Social and Clinical Psychology, 17,* 1-10. doi:10.1521/jscp.1998.17.1.1

Houlcroft, L., Bore, M., & Munro, D. (2012). Three faces of narcissism. *Personality and Individual Differences, 53,* 274–278. doi:10.1016/j.paid.2012.03.036

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. doi:10.1080/10705519909540118

Huang, G. H., & Gove, M. (2012). Confucianism and Chinese families: Values and practices in education. *International Journal of Humanities and Social Science, 2,* 10-14.

Hyler, S. E. (1994). *PDQ-4 and PDQ-4+ instructions for use*. New York, NY: New York State Psychiatric Institute.

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced issues. *Journal of Consumer Psychology, 20,* 90-98. doi:10.1016/j.jcps.2009.09.003

IBM Corp. (2013). *IBM SPSS Statistics for Windows* (*Version 22.0*) [Computer software]. Armonk, NY: IBM Corp.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38,* 1212-1218. doi:10.1111/j.1365-2929.2004.02012.x

Jonason, P. K., Li. N. P., & Czarna, A. Z. (2013). Quick and dirty: Some psychosocial costs associated with the dark triad in three countries. *Evolutionary Psychology, 11,* 172-185.

Jung, J., Forbes, G. B., & Lee, Y. (2009). Body dissatisfaction and disordered eating among early adolescents from Korea and the US. *Sex Roles, 61,* 42-54. doi:10.1007/s11199-009-9609-5

Kam, H. W., & Gopinathan, S. (1999). Recent developments in education in Singapore. *School Effectiveness and School Improvement, 10,* 99-117. doi:10.1076/sesi.10.1.99.3512

Kanfer, F. H. (1979). Personal control, social control, and altruism: Can society survive the age of individualism? *American Psychologist, 34,* 231-239. doi:10.1037//0003-066x.34.3.231

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research,* Advance online publication. doi:10.1177/0049124114543236

Kerig, P. K., & Stellwagen, K. K. (2010). Roles of callous-unemotional traits, narcissism, and machiavellianism in childhood aggression. *Journal of Psychopathology and Behavioral Assessment. 32,* 343-352. doi:10.1007/s10862-009-9168-7

Kernberg, O. (1967). Borderline personality organization. *Journal of the American Psychoanalytic Association, 15,* 641-685. doi:10.1177/000306516701500309

Kernberg, O. (1970). Factors, psychoanalytic treatment of narcissistic personalities. *Journal of the American Psychoanalytic Association, 18,* 51-85. doi:10.1177/000306517001800103

Kernberg, O. (1975). *Borderline conditions and pathological narcissism.* New York, NY: Aronson.

Kernberg, O. (1976). *Object relations theory and clinical psychoanalysis.* New York, NY: Aronson.

Khoo, M., & Karan, K. (2007). Macho or metrosexual: The branding of masculinity in FHM magazine in Singapore, *Intercultural Communication Studies, XVI,* 34-45.

Kim, K. H. (2005). Learning from each other: Creativity in East Asian and American education. *Creativity Research Journal, 17,* 337-347. doi:10.1207/s15326934crj1704_5

King, R. B., Ganotice, F. A., Jr., & Watkins, D. A. (2012). Cross-cultural validation of the Inventory of School Motivation (ISM) in the Asian setting: Hong Kong and the Philippines. *Child Indicators Research,* 5, 135-153. doi:10.1007/s12187-011-9117-3

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology, 72,* 1245-1267. doi:10.1037/0022-3514.72.6.1245

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford Press.

Klonsky, E. D., Jane, J. S., Turkheimer, E., & Oltmanns, T. F. (2002). Gender role and personality disorders. *Journal of Personality Disorders, 16,* 465-477. doi:10.1521/pedi.16.5.464.22121

Kohut, H. (1968). The psychoanalytic treatment of narcissistic personality disorders - Outline of a systematic approach. *Psychoanalytic Study of the Child, 23,* 86-113.

Kohut, H. (1971). *The analysis of the self.* New York, NY: International Universities Press.

Kohut, H. (1977). *The restoration of the self.* New York, NY: International Universities Press

Kohut, H. (1984). *How does analysis cure?* Chicago, IL: University of Chicago Press.

Kubarych, T. S., Deary, I. J., & Austin, E. J. (2004). The Narcissistic Personality Inventory: Factor structure in a non-clinical sample. *Personality and Individual Differences, 36,* 857–872. doi:10.1016/s0191-8869(03)00158-2

Kwan, V. S. Y., Kuang, L. L., & Hui, N. H. H. (2009). Identifying the sources of self-esteem: The mixed medley of benevolence, merit, and bias. *Self and Identity, 8,* 176–195. doi:10.1080/15298860802504874

Kvaraceus, W. C. (1944). Chronological ages of 761 delinquents at time of initial apprehension. *Journal of Criminal Law and Criminology, 35,* 166-168. doi:10.2307/1138755

Lapsley, D. K., Fitzgerald, D. P., Rice, K. G., & Jackson, S. (1989). Separation-individuation and the "new look" at the imaginary audience and personal fable: A test of an integrative model. *Journal of Adolescent Research, 4,* 483-505. doi:10.1177/074355488944006

Lasch, C. (1978). *The culture of narcissism: American life in an age of diminishing expectations.* New York, NY: W. W. Norton.

Levy, K. N., Ellison, W. D., & Reynoso, J. S. (2011*). A historical review of narcissism and narcissistic personality.* In W. K. Campbell & J. D. Miller (Eds.), *The handbook of narcissism and narcissistic personality disorder: Theoretical approaches, empirical findings, and treatments* (pp. 3-13). Hoboken, NJ: Wiley.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49,* 764-766. doi:10.1016/j.jesp.2013.03.013

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53–76. doi:10.1207/s15327906mbr3201_3

Livesley, W. J., Jang, K. L., Jackson, D. N., & Vernon, P. A. (1993). Genetic and environmental contributions to dimensions of personality disorder. *American Journal of Psychiatry, 150,* 12-20.

Loke, S. W., & Lowe, P. A. (2014). Validation of the Narcissistic Personality Questionnaire for Children—Revised among U.S. students. *Psychological Assessment, 26,* 619-627. doi:10.1037/a0035442

Lowe, P. A. (2014). Should test anxiety be measured differently for males and females? Examination of measurement bias across gender on measures of test anxiety for middle and high school, and college students. *Journal of Psychoeducational Assessment,* Advance online publication. doi:10.1177/0734282914549428

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130-149. doi:10.1037/1082-989x.1.2.130

Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69,* 1030-1053. doi:10.1111/j.1467-8624.1998.tb06159.x

McIntire, S. A., & Miller, L. A. (2007). *Foundations of psychological testing: A practical approach* (2nd ed.). Thousand Oaks, CA: Sage.

Melville, A. D. (1986). *Ovid metamorphoses: Oxford word's classics.* New York, NY: Oxford.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58,* 525-542. doi:10.1007/bf02294825

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research, 3,* 111-121.

Miller, J. D., & Campbell, W. K. (2011). Addressing criticisms of the Narcissistic Personality Inventory (NPI). In W. K. Campbell & J. D. Miller (Eds.), *Handbook of narcissism and narcissistic personality disorder: Theoretical approaches, empirical findings, and treatments* (pp. 146-152). Hoboken, NJ: Wiley.

Miller, J. D., Hoffman, B. J., Campbell, W. K., & Pilkonis, P.A. (2008). An examination of the factor structure of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, narcissistic personality disorder criteria: One or two factors. *Comprehensive Psychiatry, 49,* 180-192. doi:10.1016/j.comppsych.2007.08.01

Miller, J. D., Hoffman, B. J., Gaughan, E. T., Gentile, B., Maples, J., & Campbell, W. K. (2011). Grandiose and vulnerable narcissism: A nomological network analysis. *Journal of Personality, 79,* 1013-1042. doi:10.1111/j.1467-6494.2010.00711.x

Millon, T. (1981). *Disorders of personality: DSM-III, Axis II*. New York, NY: Wiley.

Millon, T. (1982). *Millon Clinical Multiaxial Inventory* (2nd ed.). Minneapolis, MN: Interpretive Scoring System.

Millon, T., Millon, C. M., Meagher, S., Grossman, S., & Ramnath, R. (2004). *Personality disorders in modern life.* Hoboken, NJ: John Wiley & Sons.

Ministry of Education, Singapore. (2014). *Subject-based banding in primary schools: Catering to your child's abilities*. Retrieved from http://www.moe.gov.sg/education/primary/files/subject-based-banding.pdf

Morey, L. C., Waugh, M. H., & Blashfield, R. K. (1985). MMPI scales for DSM-III personality disorders: Their derivation and correlates. *Journal of Personality Assessment, 49,* 245-251. doi:10.1207/s15327752jpa4903_5

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry, 12,* 177-196. doi:10.1207/s15327965pli1204_1

Muthén, L. K., & Muthén, B. O. (1998-2013), *MPlus user's guide (7th ed.)*. Los Angeles, CA: Muthén & Muthén.

Nicholls, J. G. (1990). What is ability and why are we mindful of it? A developmental perspective. In R. J. Sternberg & J. Kolligian, Jr. (Eds.), *Competence considered* (pp. 11–40). New Haven, CT: Yale University Press.

Ojanen, T., Findley, D., & Fuller, S. (2012) Physical and relational aggression in early adolescence: Associations with narcissism, temperament, and social goals. *Aggressive Behavior, 38,* 99-107. doi:10.1002/ab.21413

Oswin, N., & Yeoh, B. S. A. (2010). Introduction: Mobile city Singapore. *Mobilities, 5,* 167-175. doi:10.1080/17450101003665002

Ow, R. (2014). Social work in Singapore. In C. Aspalter. (Ed.), *Social work in East Asia* (pp. 123-138). Burlington, VA: Ashgate.

Pepper, L. J., & Strong, P. N. (1958). Judgmental subscales for the MF scale of the MMPL Unpublished manuscript.

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N. (2009). Initial construction and validation of the Pathological Narcissism Inventory. *Psychological Assessment, 21,* 365-379. doi:10.1037/a0016530

Pincus, A. L., & Lukowitsky, M. R. (2010). Pathological narcissism and narcissistic personality disorder. *Annual Review of Clinical Psychology, 6,* 421-446. doi:10.1146/annurev.clinpsy.121208.131215

Plakun, E. M. (1989). Narcissistic personality disorder: A validity study and comparison to borderline personality disorder. *Psychiatric Clinics of North America, 12,* 603–620.

Pottthoff, R. F. (1966). *Statistical aspects of the problem of biases in psychological tests.* Institute of Statistics Mimeo Series No. 479. Chapel Hill, NC: University of North Carolina.

Raskin, R. & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of Personality and Social Psychology, 54,* 890-902. Doi:10.1037//0022-3514.54.5.890

Raskin, R. N., & Hall, C. S. (1979). A Narcissistic Personality Inventory. *Psychological Reports, 45*, 590. doi:10.1207/s15327752jpa4502_10

Rathvon, N., & Holmstrom, R. W. (1996). An MMPI-2 portrait of narcissism. *Journal of Personality Assessment, 66,* 1-19. doi:10.1207/s15327752jpa6601_1

Reynolds, C. R. (1982). Methods for detecting construct and predictive bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias.* Baltimore, MD: John Hopkins University Press.

Reynolds, C. R., & Lowe, P. A. (2009). The problem of bias in psychological assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (4th ed., pp. 332-374). Hoboken, NJ: John Wiley & Sons.

Reynolds, C. R., Lowe, P. A., & Saenz, A. (1999). The problem of bias in psychological assessment. In T. B. Gutkin & C. R. Reynolds (Eds.), *The handbook of school psychology* (3rd ed., 549–595). New York, NY: Wiley.

Reynolds, C. R., & Suzuki, L. A. (2013). Bias in psychological assessment: An empirical review and recommendation. In I. B. Weiner, J. R. Graham, & J. A. Naglieri (Eds.), *Handbook of Psychology: Vol. 10. Assessment psychology* (2nd ed., pp. 82-113). Hoboken, NJ: Wiley.

Reynolds, M. R., & Keith, T. Z. (2013). Measurement and statistical issues in child assessment research. In D. H. Soklofske, C. R. Reynolds, & V. Schwean (Eds.), *Oxford handbook of child and adolescent assessment* (pp.48-83). New York, NY: Oxford University Press.

Ronningstam, E. F. (2005). *Identifying and understanding the narcissistic personality.* New York, NY: Oxford University Press.

Rose, P. (2002). The happy and unhappy faces of narcissism. *Personality and Individual Differences, 33,* 379-391. doi:10.1016/s0191-8869(01)00162-3

Rudman, L. A., Moss-Racusin, C. A., Glick, P., & Phelan, J. E. (2012). Reactions to vanguards: Advances in backlash theory. In P. G. Devine & E. A. Plant (Eds.), *Advances in Experimental Social Psychology: Vol. 45* (pp. 167-227). Burlington, MA: Elsevier.

Russell, G. A. (1985). Narcissism and the narcissistic personality disorder: A comparison of the theories of Kernberg and Kohut. *British Journal of Medical Psychology, 58,* 137-148. doi:10.1111/j.2044-8341.1985.tb02626.x

Salekin, R. T. (2006). Psychopathy in children and adolescents: Key issues in conceptualization and assessment. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 389-414). New York, NY: Guilford Press.

Salmivalli, C. (2001). Feeling good about oneself, being bad to others? Remarks on self-esteem, hostility, and aggressive behavior. *Aggression and Violent Behavior, 6,* 375-393. doi:10.1016/s1359-1789(00)00012-4

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 24,* 347-363. doi:10.1177/0734282911406661

Seah, S. L., & Ang, R. P. (2008). Differential correlates of reactive and proactive aggression in Asian adolescents: Relations to narcissism, anxiety, schizotypal traits, and peer relations. *Aggressive Behavior, 34,* 553-562. doi:10.1002/ab.20269

Sedikides, C., Rudich, E. A., Gregg, A. P., Kumashiro, M., & Rusbult, C. (2004). Are normal narcissists psychologically healthy?: Self-esteem matters. *Personality Processes and Individual Differences, 87,* 400-416. doi:10.1037/0022-3514.87.3.400

Serkownek, K. (1975). *Subscales for scale 5 and 0 of the MMPI.* Unpublished manuscript.

Sharma, G. (2004). A comparative study of the personality characteristics of primary-school students with learning disabilities and their nonlearning disabled peers. *Learning Disabilities Quarterly, 27,* 127-140. doi:10.2307/1593664

Singapore Department of Statistics. (2014). *Population trends 2014.* Retrieved from http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/population_and_population_structure/population2014.pdf

Solano-Flores, G. (2011). Assessing the cultural validity of assessment practices: An introduction. In M. R. Basterra, E. Trumbull, & G. Solano (Eds.), *Cultural validity in assessment: A guide for educators* (pp. 3-21). New York, NY: Routledge.

Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching, 38,* 553–573. doi: dx.doi.org/10.1002/tea.1018

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2009). Testing measurement invariance using multigroup CFA: Differences between educational groups in human values measurement. *Quality and Quantity, 43,* 599-616. doi:10.1007/s11135-007-9143-x

Tamborski, M., & Brown, R. P. (2011). The measurement of trait narcissism in social-personality research. In W. K. Campbell & J. D. Miller (Eds.), *The handbook of narcissism and narcissistic personality disorder: Theoretical approaches, empirical findings, and treatment.* Hoboken, NJ: John Wiley & Sons.

Tamney, J. B. (1995). *The struggle over Singapore's soul: Western modernization and Asian culture*. New York, NY: de Gruyter.

Teh, L. W. (2014). Singapore's performance in PISA: Levelling up the long trail. In S. K. Lee, W. O. Lee, & E. L. Low (Eds.), *Educational policy innovations: Levelling up and sustaining educational achievement* (pp. 71-84). Singapore: Springer.

The World Bank Group. (2015). *GDP per capita* (*current US$*). Retrieved from http://data.worldbank.org/indicator/NY.GDP.PCAP.CD

Thomaes, S., Bushman, B. J., Orobio de Castro, B., & Stegge, H. (2009). What makes narcissists bloom? A framework for research on the etiology and development of narcissism. *Development and Psychopathology, 21,* 1233-1247. doi:10.1017/s0954579409990137

Thomaes, S., Stegge, H., Bushman, B. J., Olthof, T., & Denissen, J. (2008). Development and validation of the Childhood Narcissism Scale. *Journal of Personality Assessment, 90,* 382–391. doi:10.1080/00223890802108162

Thompson, C., & Hirschman, E. (1995). Understanding the socialized body: A poststructuralist analysis of consumers' self-conceptions, body images, and self-care practices. *Journal of Consumer Research, 22,* 139−153. doi:10.1086/209441

Tracy, J. L., & Robins, R. W. (2004a). Putting the self into self-conscious emotions: A theoretical model. *Psychological Inquiry, 15,* 103–125. doi:10.1207/s15327965pli1502_01

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review, 96,* 506-520. doi:10.1037//0033-295x.96.3.506

Triandis, H. C. (2001). Individualism-collectivism and personality. *Journal of Personality, 69,* 907–924. doi:10.1111/1467-6494.696169

Tschanz, B. T., Morf, C. C., & Turner, C. W. (1998). Gender differences in the structure of narcissism: A multi-sample analysis of the Narcissistic Personality Inventory. *Sex Roles, 38,* 863–870. doi:10.1023/a:1018833400411

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10. doi:10.1007/bf02291170

Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles, 36,* 305–325. doi:10.1007/bf02766650

Twenge, J. M. (2006). *Generation me: Why today's young Americans are more confident, assertive, entitled – and more miserable than ever before.* New York, NY: Free Press.

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *Journal of Personality, 76,* 875-902. doi:10.1111/j.1467-6494.2008.00507.x

van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 54,* 119-135. doi:10.1016/j.erap.2003.12.004

Vernon, P. A., Villani, V. C., Vickers, L. C., & Harris, J. A. (2008). A behavioural genetic investigation of the Dark Triad and the Big 5. *Personality and Individual Differences, 44,* 445–452. doi:10.1016/j.paid.2007.09.007

Vitacco, M. J., Rogers, R., & Neumann, C. S. (2003). The Antisocial Processing Screening Device: An examination of its construct and criterion validity. *Assessment, 10,* 143-50. doi:10.1177/1073191103010002005

Vogel, D. L., Wade, N. G., & Haake, S. (2006). Measuring the self-stigma associated with seeking psychological help. *Journal of Counseling Psychology, 53,* 325–337. doi:10.1037/0022-0167.53.3.325

Washburn, J. J., McMahon, S. D., King, C. A., Reinecke, M. A., & Silver, C. (2004). Narcissistic features in young adolescents: Relations to aggression and internalizing symptoms. *Journal of Youth and Adolescence, 33,* 247–260. doi:10.1023/b:joyo.0000025323.94929.d9

Watson, P. J., Little, T., Sawrie, S. M., & Biderman, M. D. (1992). Measures of the narcissistic personality: Complexity of relationships with self-esteem and empathy. *Journal of Personality Disorders, 6,* 433-448. doi:10.1521/pedi.1992.6.4.434

Wilson, M. S., & Sibley, C. G. (2011). 'Narcissism creep': Evidence for age-related differences in narcissism in the New Zealand general population. *New Zealand Journal of Psychology, 40,* 87-93.

Wink, P. (1991). Two faces of narcissism. *Journal of Personality and Social Psychology, 61,* 590–597. doi:10.1037//0022-3514.61.4.590

Wink, P. (1996). Narcissism. In C. G. Costello (Ed.), *Personality characteristics of the personality disordered* (pp. 146-172). New York, NY: Wiley.

Wink, P., & Gough, H. G. (1990). New narcissism scales for the California Psychological Inventory and MMPI. *Journal of Personality Assessment, 54,* 446-463. doi:10.1080/00223891.1990.9674010

Wirth, L. (2001). *Breaking through the glass ceiling: Women in management.* Geneva, Switzerland: International Labour Office.

Wong, F. M. (1981). Effects of the employment of mothers on marital role and power differentiation in Hong Kong. In A. Y. C. King & R. P. L. Lee (Eds.), *Social Life and Development in Hong Kong* (pp. 217-234). Hong Kong: Chinese University Press.

Wright, A. G. C., Lukowitsky, M. R., Pincus, A. L., & Conroy, D. E. (2010). The higher order factor structure and gender invariance of the Pathological Narcissism Inventory. *Assessment, 17,* 467-483. doi:10.1177/1073191110373227

Yang, C. J., Gray, P., & Pope, H. G. Jr. (2005). Male body image in Taiwan versus the West: Yanggang Zhiqi meets the Adonis complex. *American Journal of Psychiatry, 162,* 263-269. doi:10.1176/appi.ajp.162.2.263

Zhou, H., Li, Y., Zhang, B., & Zeng, M. (2012). The relationship between narcissism and friendship qualities in adolescents: Gender as a moderator. *Sex Roles, 67,* 452-462. doi:10.1007/s11199-012-0169-8

APPENDIX A

HSC-L Letter of Approval

KU RESEARCH & GRADUATE STUDIES
The University of Kansas

APPROVAL OF PROTOCOL

May 9, 2014

Wei Loke
swloke@ku.edu

Dear Wei Loke:

On 5/9/2014, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	Examination of construct bias in the Narcissistic Personality Questionnaire for Children - Revised (NPQC-R) across gender and culture
Investigator:	Wei Loke
IRB ID:	STUDY00001182
Funding:	None
Grant ID:	None

The IRB approved the study on 5/9/2014.

1. Any significant change to the protocol requires a modification approval prior to altering the project.
2. Notify HSCL about any new investigators not named in original application. Note that new investigators must take the online tutorial at https://rgs.drupal.ku.edu/human_subjects_compliance_training.
3. Any injury to a subject because of the research procedure must be reported immediately.
4. When signed consent documents are required, the primary investigator must retain the signed consent documents for at least three years past completion of the research activity.

Please note university data security and handling requirements for your project:
https://documents.ku.edu/policies/IT/DataClassificationandHandlingProceduresGuide.htm

You must use the final, watermarked version of the consent form, available under the "Documents" tab in eCompliance.

Sincerely,

Stephanie Dyson Elms, MPA
IRB Administrator, KU Lawrence Campus

Human Subjects Committee Lawrence
Youngberg Hall I 2385 Irving Hill Road I Lawrence, KS 66045 I (785) 864-7429 I HSCL@ku.edu I research.ku.edu

APPENDIX B

Permission Letter to Include NPQC-R Items in Dissertation

NIE
NATIONAL INSTITUTE OF EDUCATION
SINGAPORE

www.nie.edu.sg
1 Nanyang Walk Singapore 637616 Tel: (65) 6790 3888
NTU Reg No. 2006043938

April 26, 2015

Stephen Loke
University of Kansas
Department of Psychology and Research in Education
Joseph R. Pearson Hall, Rm. 621
1122 West Campus Rd.
Lawrence, Kansas 66045-3101
USA

Rebecca P. Ang, Ph.D.
Associate Professor and Head
Psychological Studies Academic Group
National Institute of Education
Nanyang Technological University
1 Nanyang Walk
Singapore 637616
SINGAPORE

Permission to include NPQC-R items in dissertation text and appendix

Dear Stephen,

I understand that you would like to include actual NPQC-R items within your dissertation. I have no objection to this, and you have my permission to include actual NPQC-R items in the text of your dissertation as well as in the appendix.

If you need further information or clarification, please do not hesitate to contact me. I can be reached by email at: rebecca.ang@nie.edu.sg

Thank you.

Yours Sincerely,

Rebecca P. Ang, Ph.D.